



Metabarcoding and Environmental DNA Workshop:

Assessing Biodiversity of a Local Pond using DNA Metabarcoding



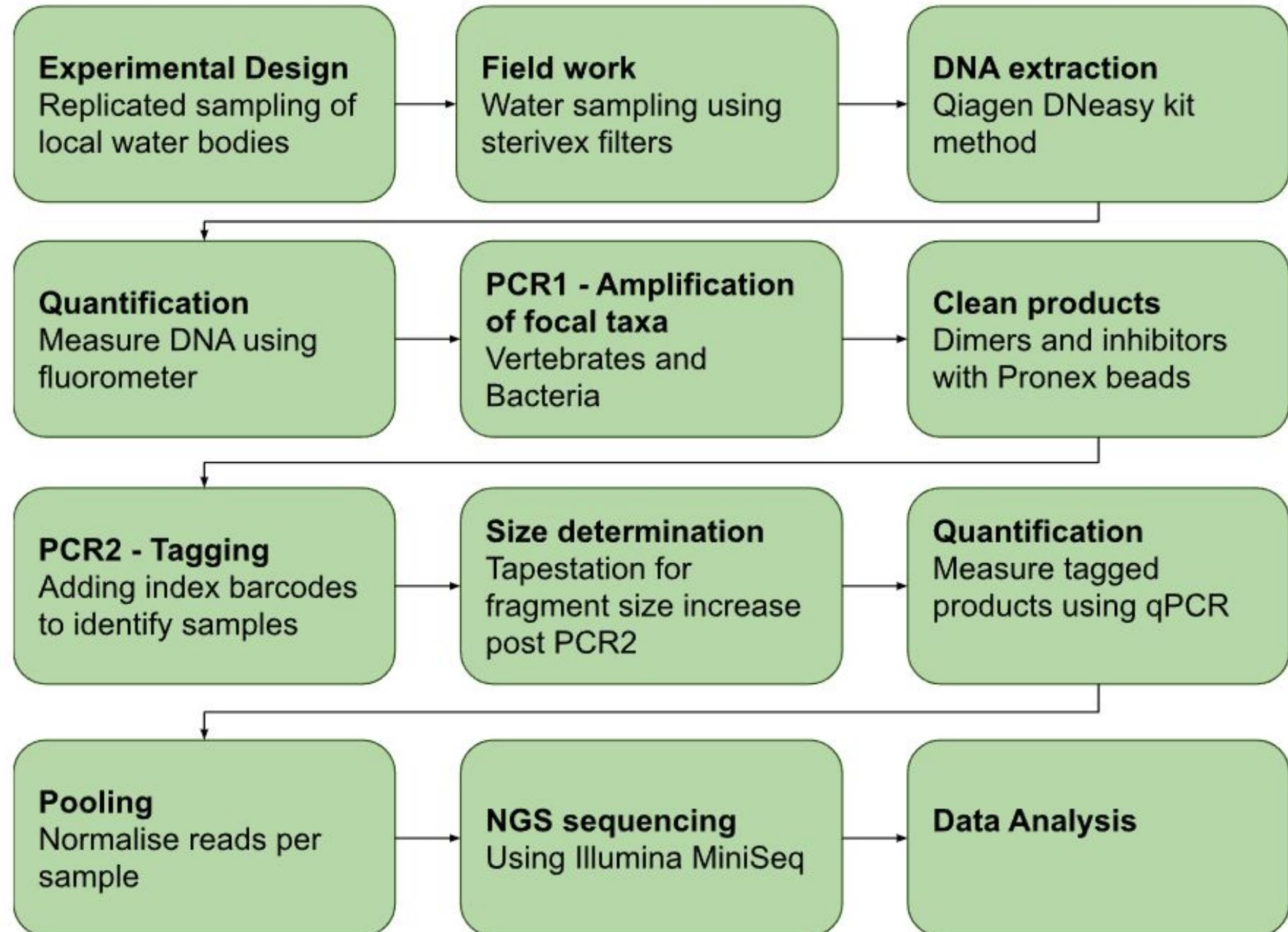
The University
Of Sheffield.



Natural
Environment
Research Council

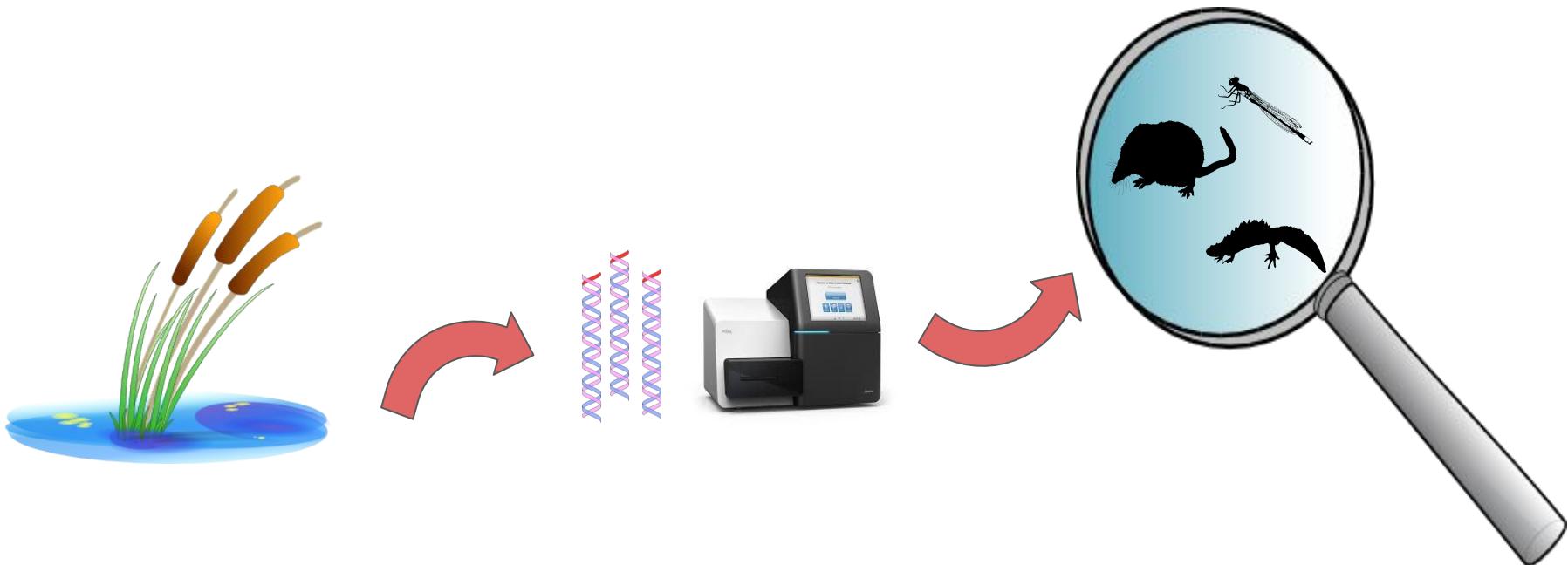


NEOF
NERC ENVIRONMENTAL
OMICS FACILITY



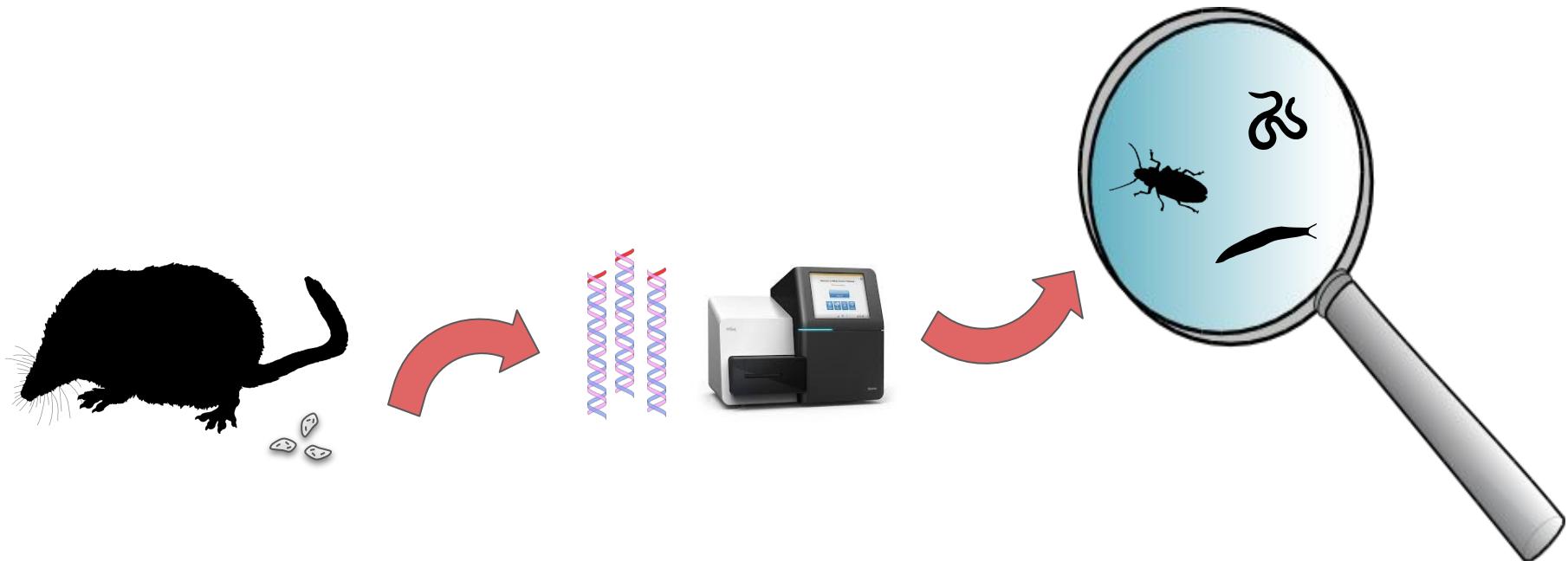
Why metabarcoding?

Identify a range of species in an environmental sample simultaneously, through genetic sequencing.



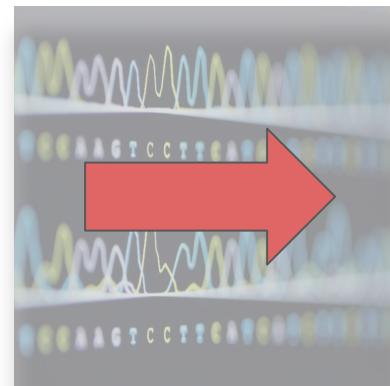
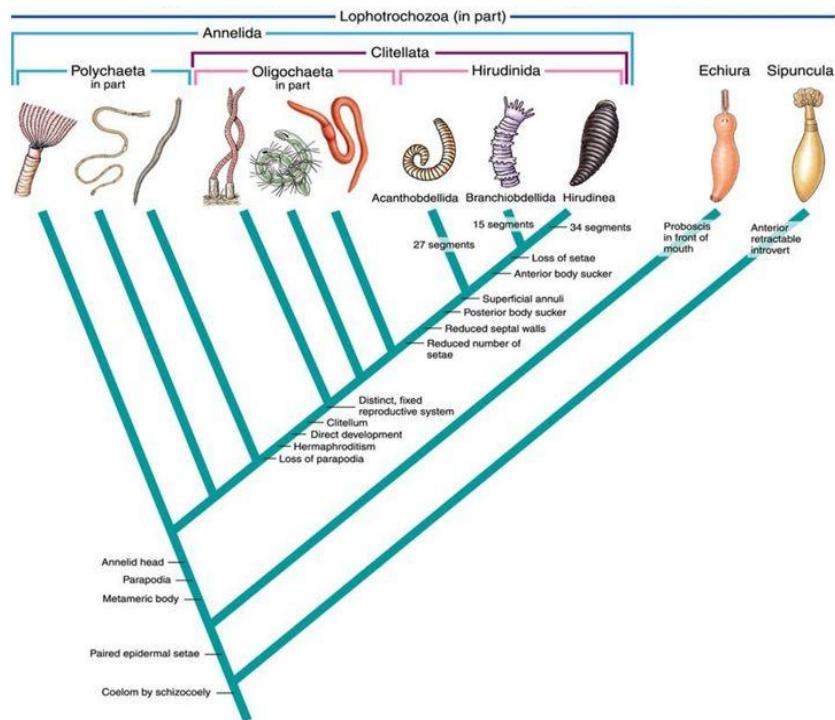
Why metabarcoding?

This allows scientists to carry out a range of studies including diet analysis and population surveys



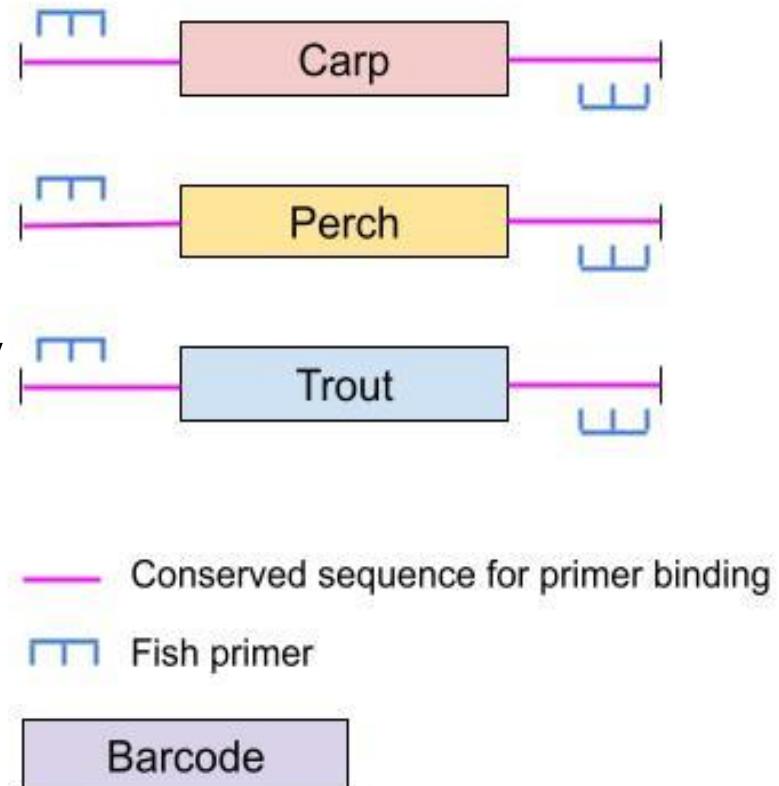
How does metabarcoding work?

Generic primers are used to target whole taxonomic groups, whilst bioinformatic analysis allows species identification

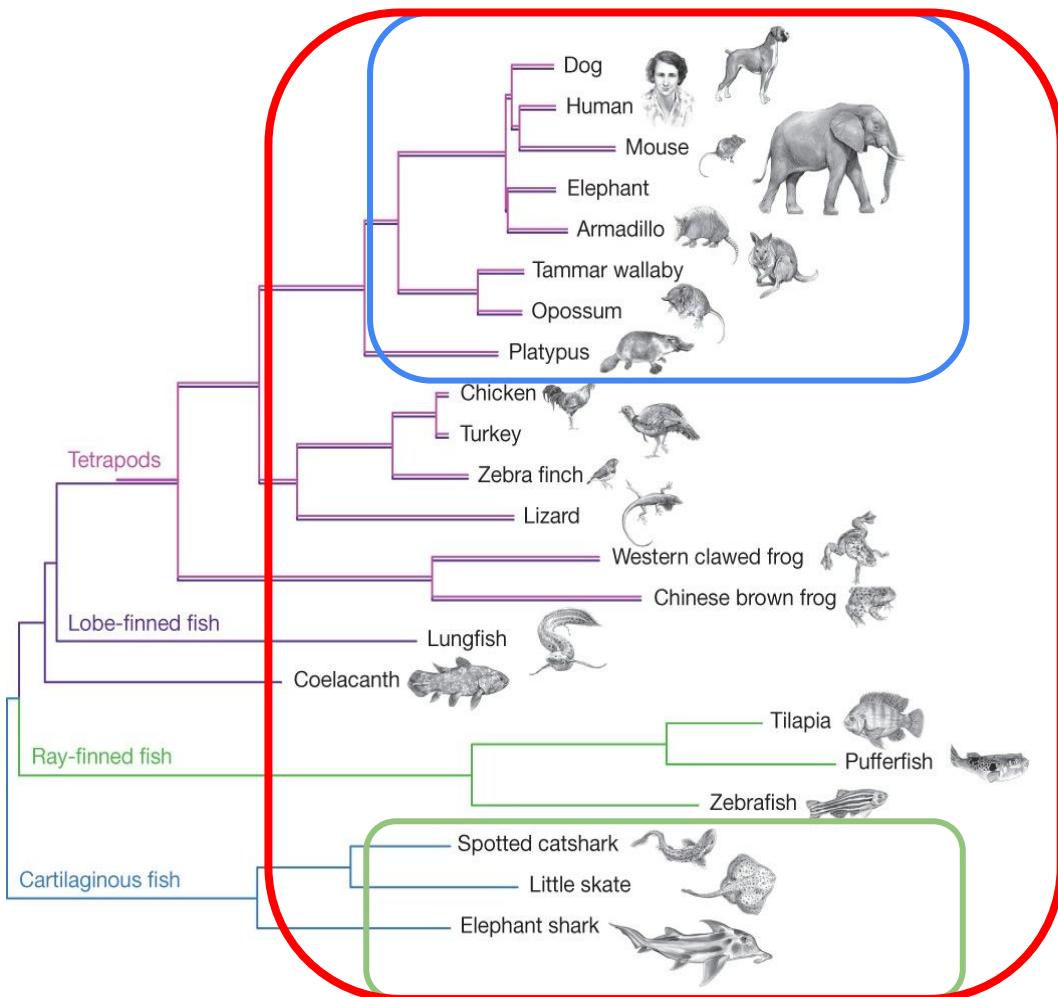


Metabarcoding genes allow species identification

- Barcoding genes are highly conserved within a species, but variable between species
- The barcoding gene should be flanked by sequences that are conserved within the target phylogeny
- Primers are designed to target the conserved sequences
- Resulting sequence data can be compared to existing barcode libraries for species identification



PCR 1 - Primer Selection



When choosing (or designing) primers it's important to have a clear focus on the scale of the phylogeny that is required to answer your question

The Diversity vs. Precision trade off

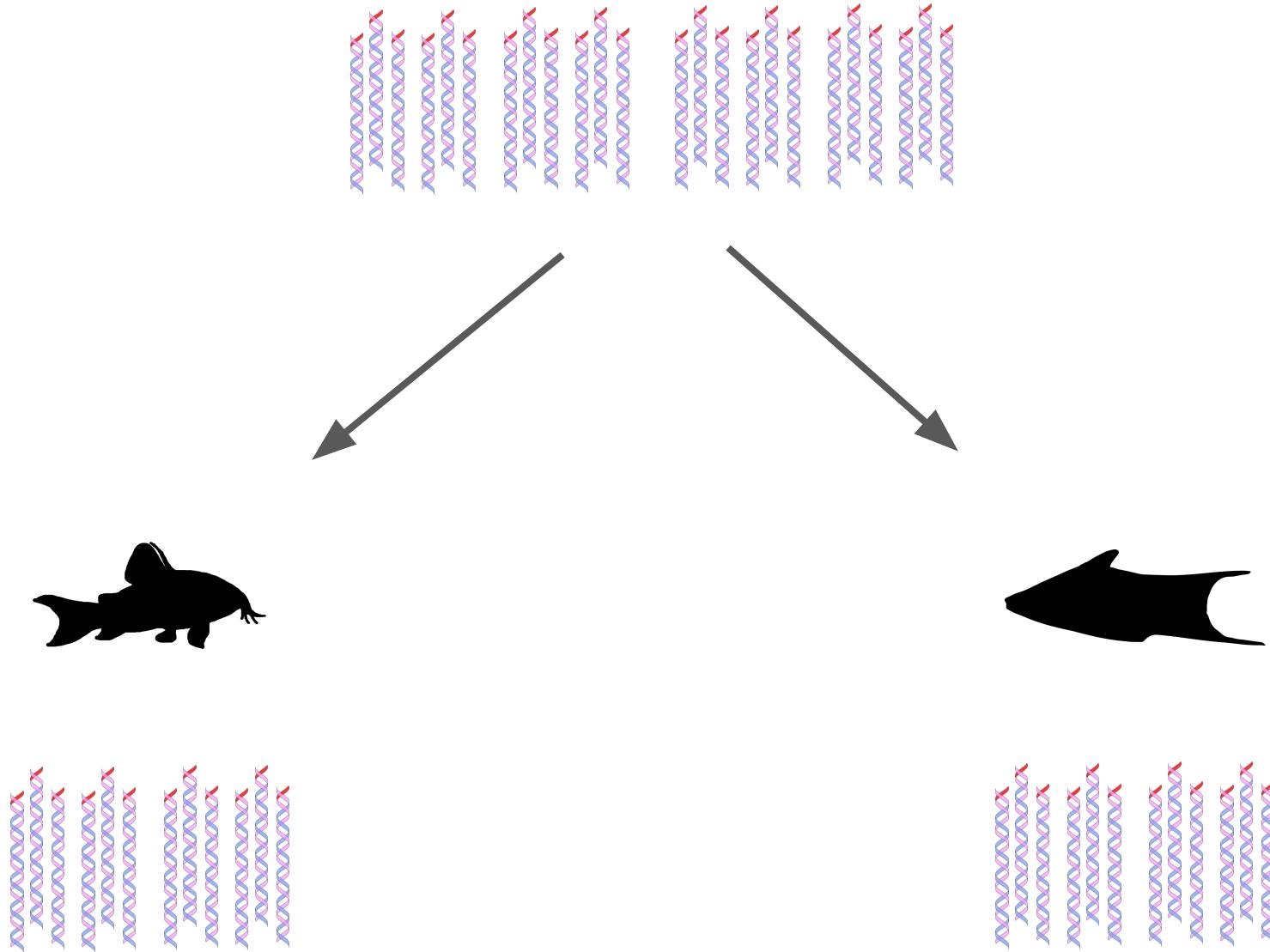
The more conserved the target region is across a phylogeny, the greater the number of potential species but the less read depth (and therefore accuracy) for each you will yield

LCO1490

MiMammal-U

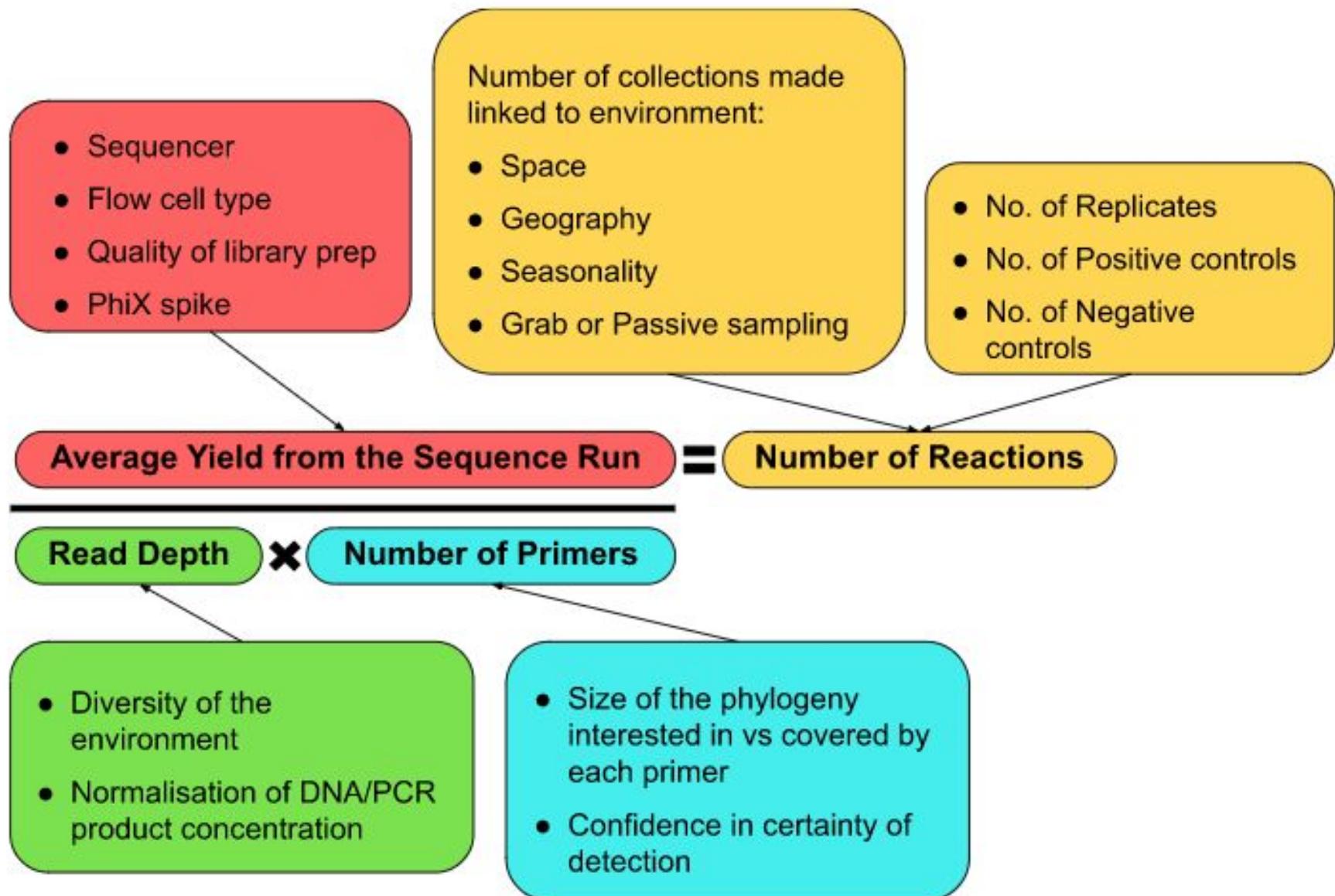
Elas02

Total number of sequencing reads to be shared between target species





Experimental Design



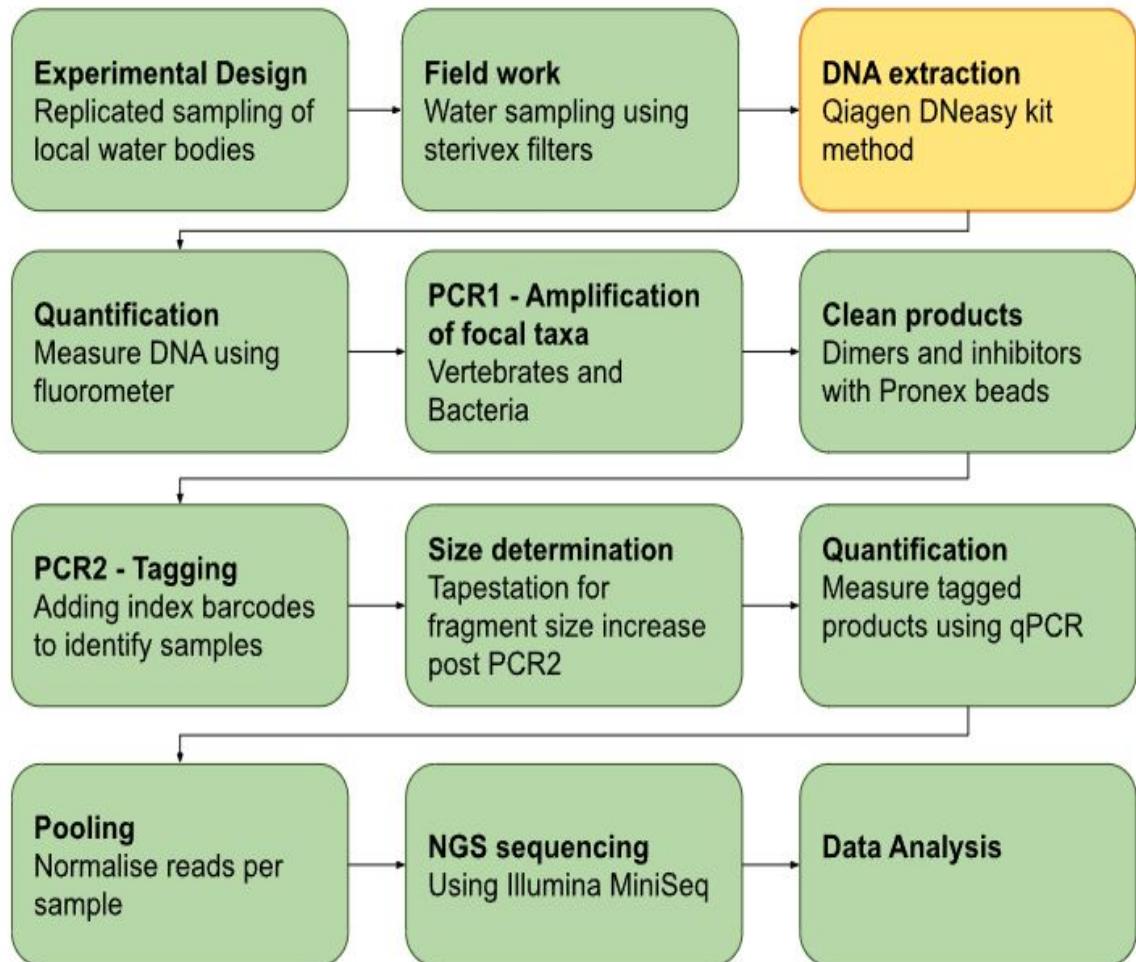
Sample Collection: Field Equipment

- Dipper
- Sterivex filter unit
- Syringe
- 250ml bottle
- Field negative bottle
- Caulking gun
- Box of ice
- 10% Bleach
- Tissue paper roll
- Gloves
- Permanent marker



DNA Extraction

Releasing and purifying DNA from the sample



DNA Extraction

- Time to start the extractions
- Firstly, for those who have had little or no experience of using a pipette let's have a quick tutorial

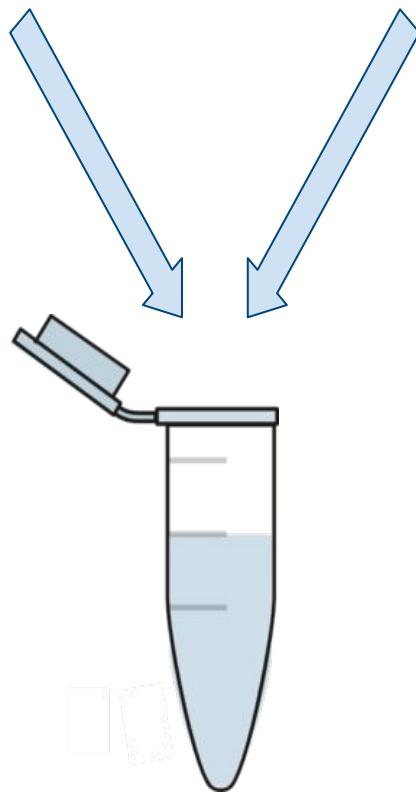




DNA Extraction: Set-up

400 µl Buffer ATL

10 µl Proteinase K



Remember to label your tubes!

DNA Extraction: Remove Filter Paper from Unit



DNA Extraction: Overnight Incubation



Incubate the samples at 56°C overnight

DNA Extraction: Qiagen DNA Purification Kit



Vortex

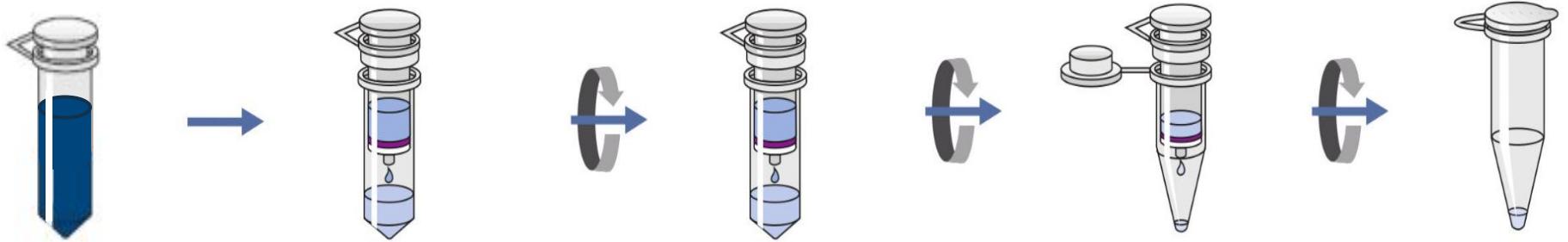


Add lysis solution to new tube



Add buffer AL and
incubate at 56°C for 10
min

Load sample
onto spin column

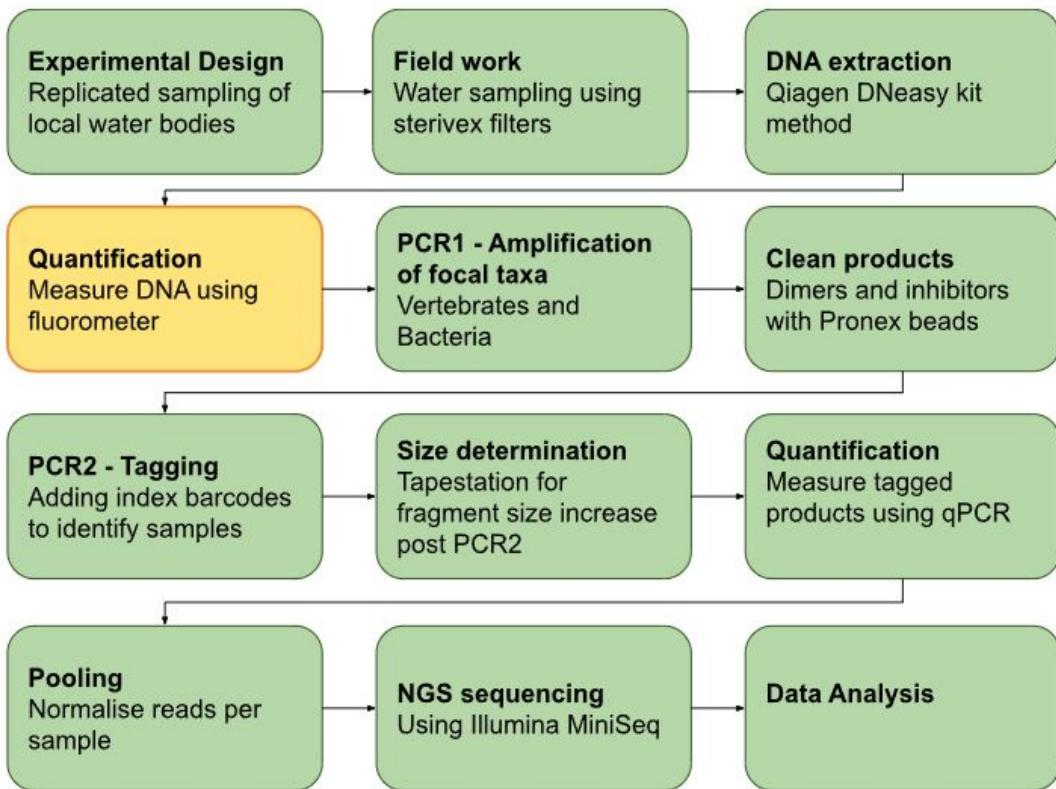


Add 100%
ethanol and
vortex

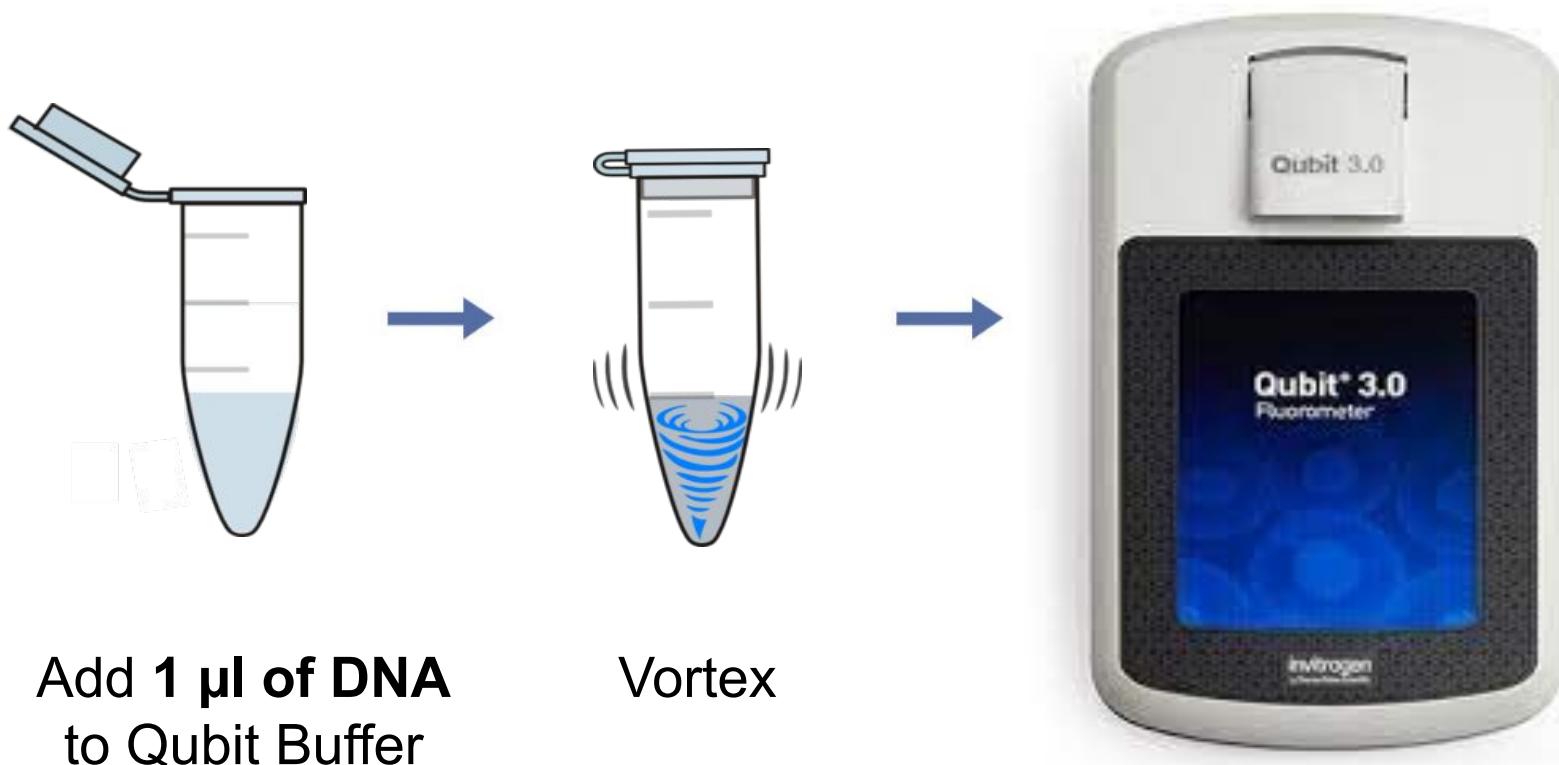
Add elution
buffer

Clean
DNA

Quantification

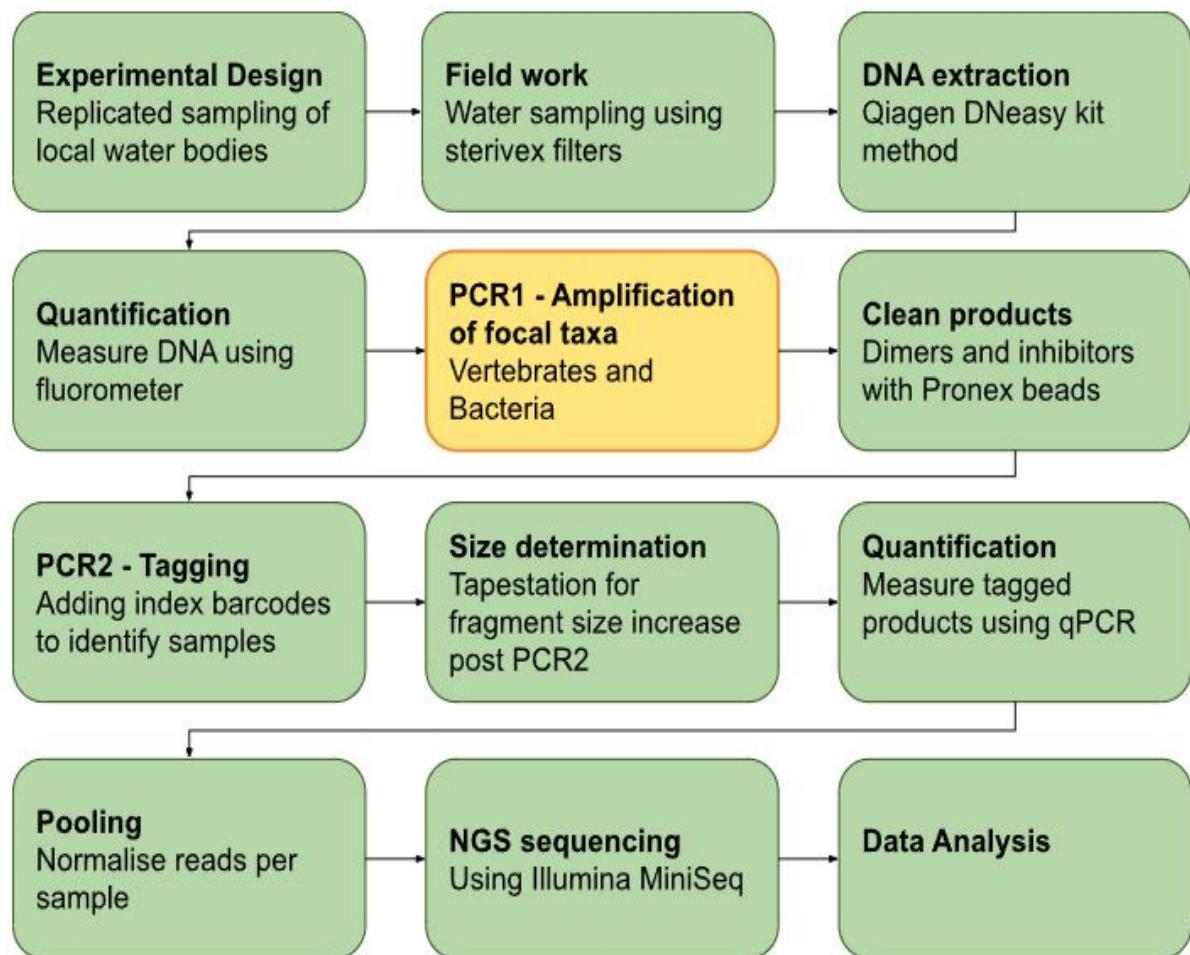


Quantification of DNA Extractions

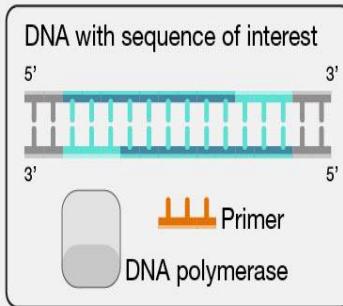


PCR 1

Amplifying your sequence of interest, and visualise with gel electrophoresis

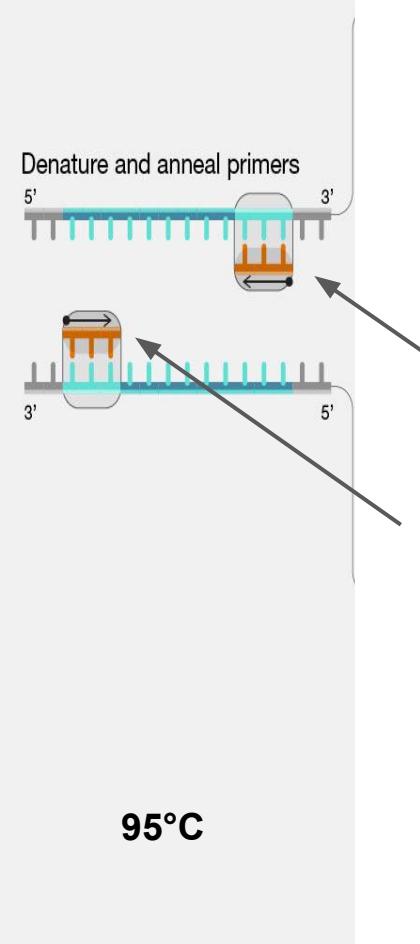


Polymerase Chain Reaction (PCR)



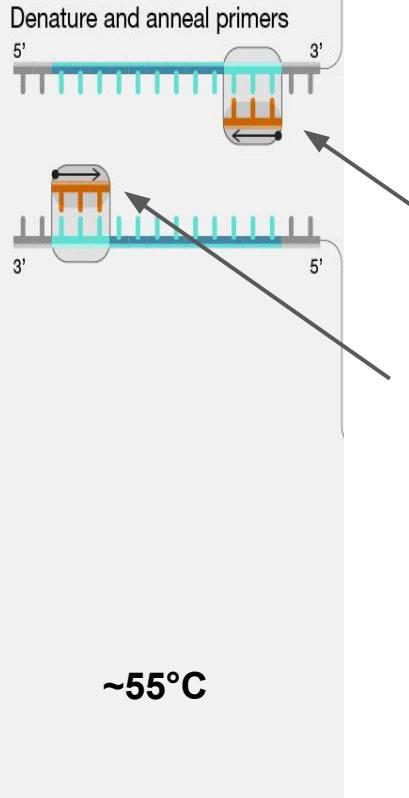
Step 1 - Denaturation

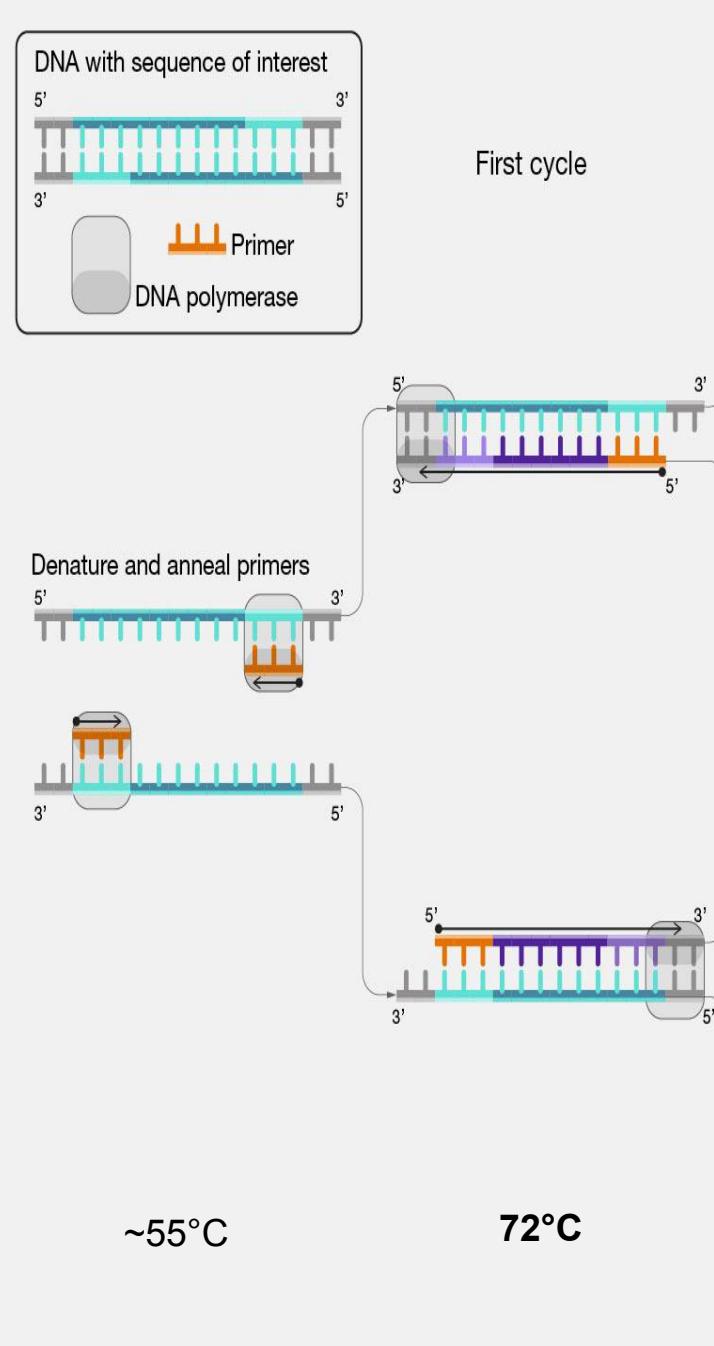
The double-stranded DNA is separated into single strands when heated



Step 2 - Annealing

- Primers bind
- Determine melting temperature

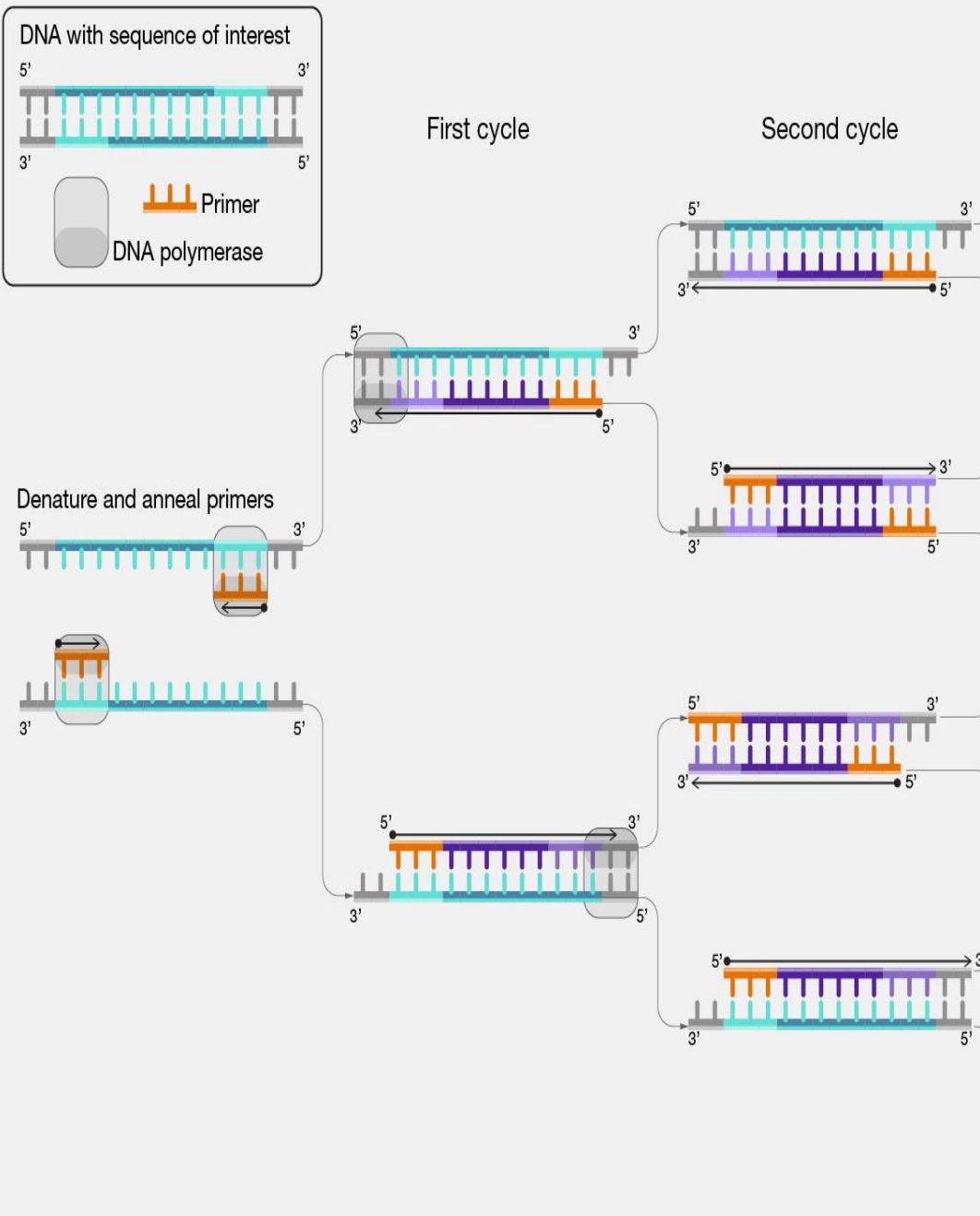




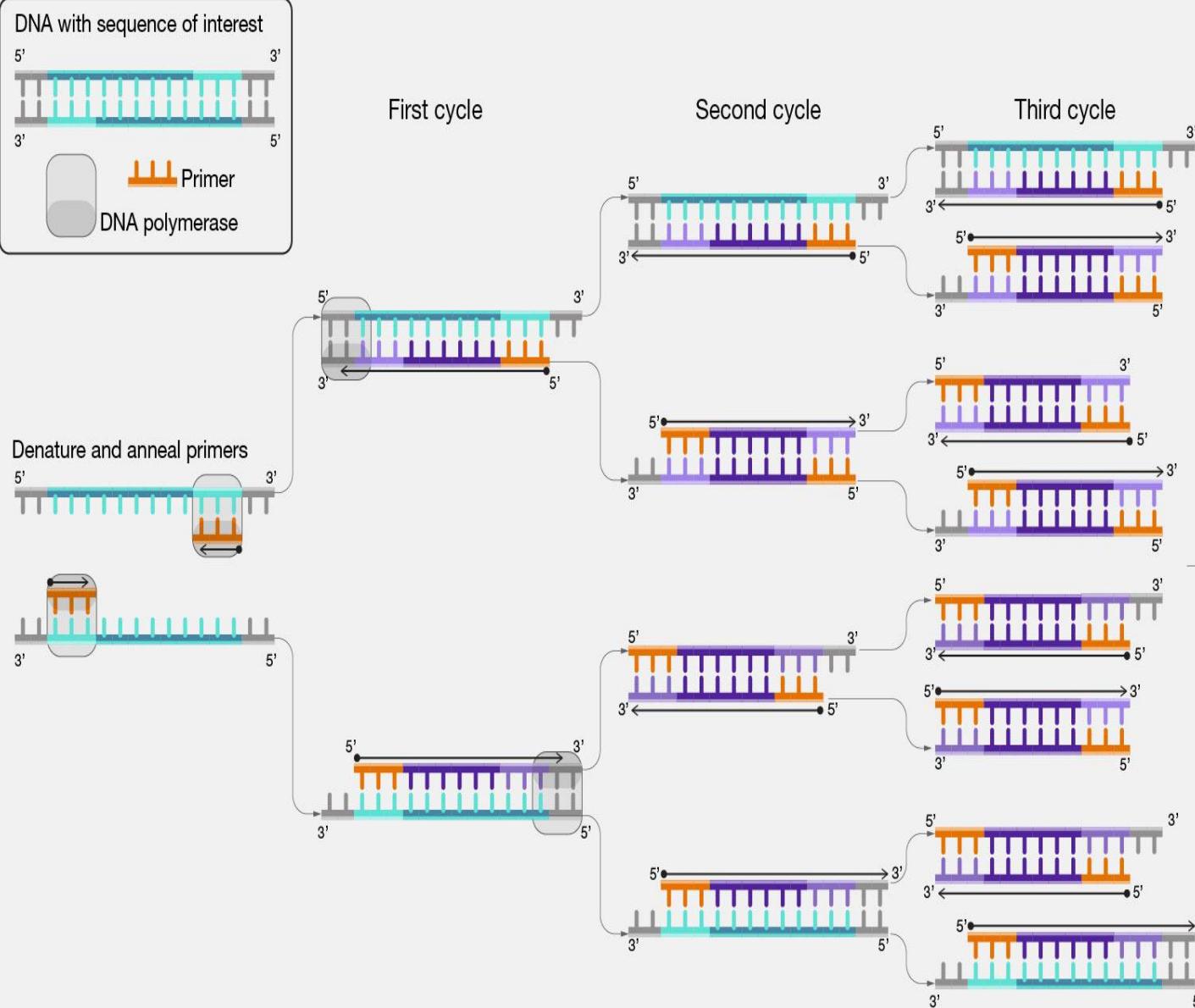
Step 3 - Extension

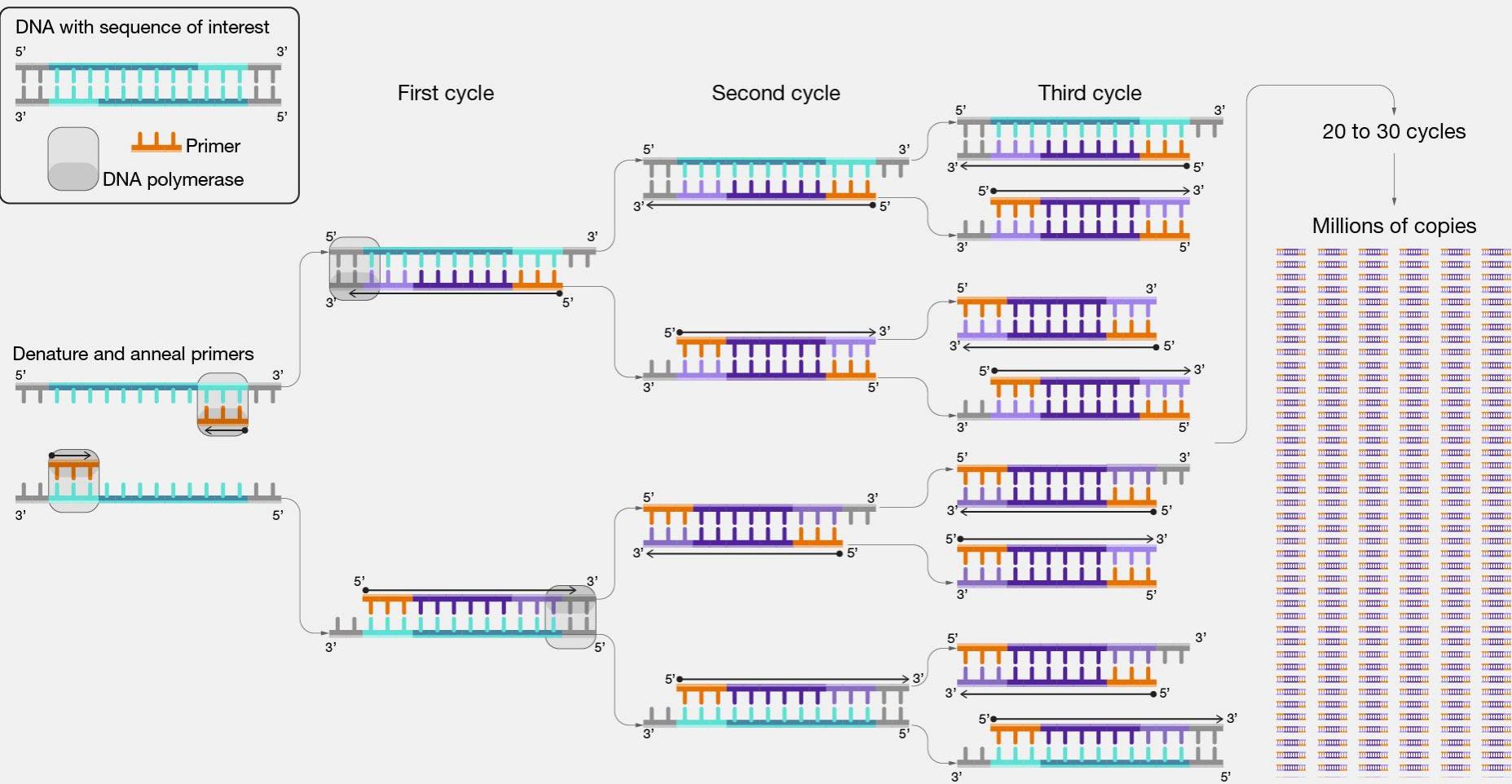
- Temperature increases
- Polymerase - amplicon extension
- Extension time

The PCR process then replicates this sequence repeatedly over 20 to 30 cycles



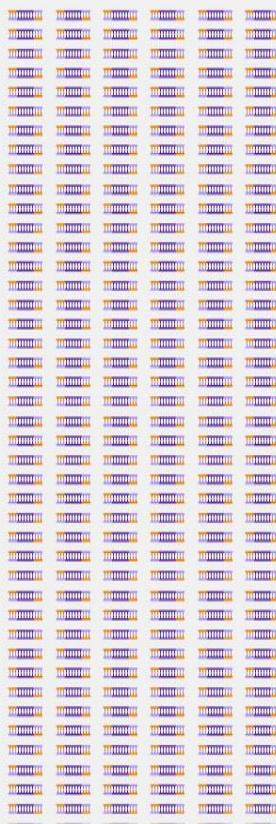
The number
of replicates
grows
exponentially





20 to 30 cycles

Millions of copies



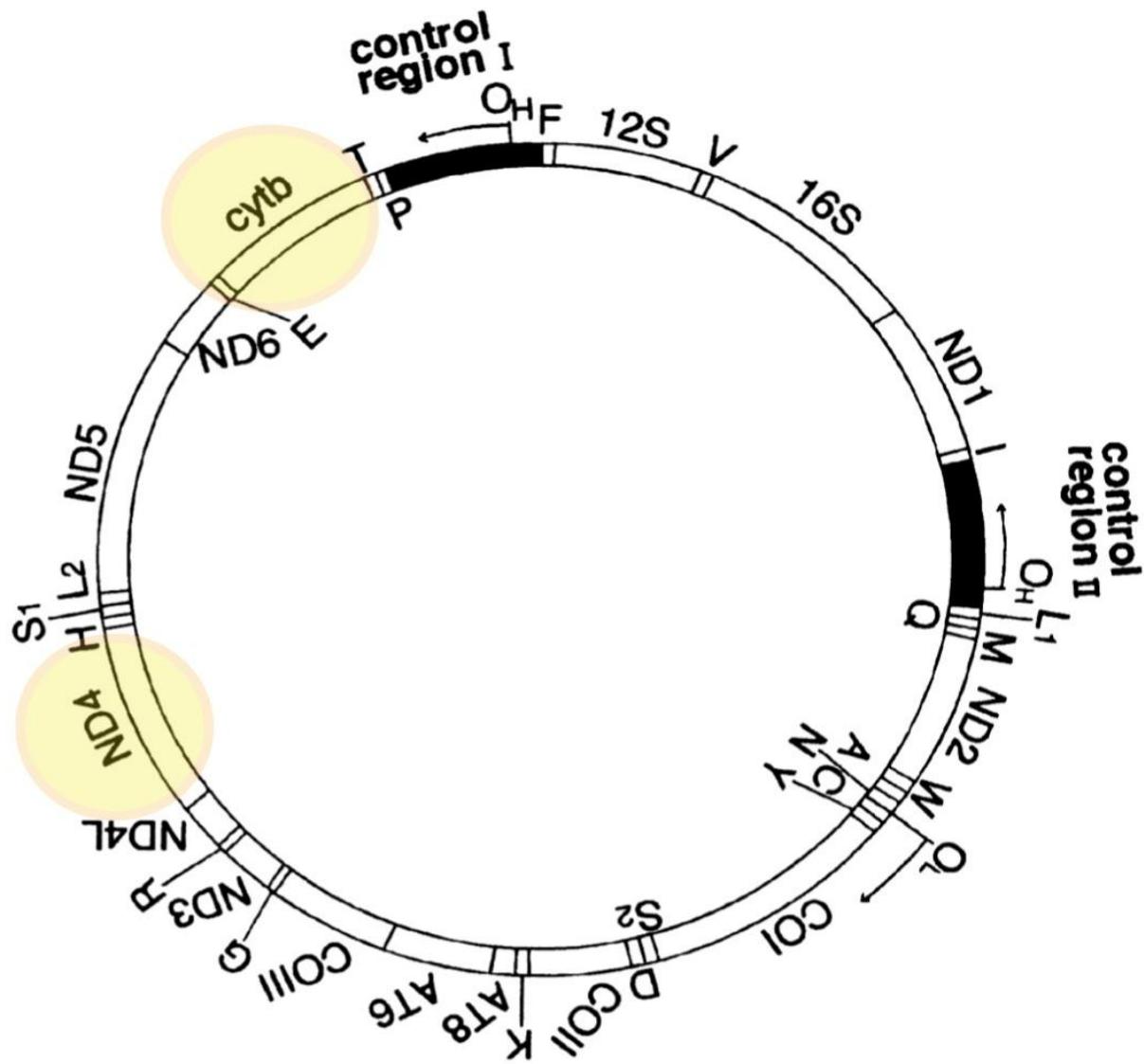
The result - millions of exact copies of your genetic sequence, known as amplicons

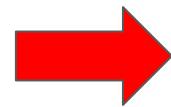


The Thermocycler

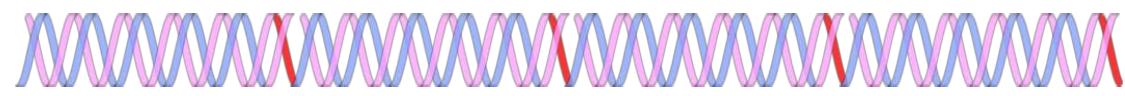
- Easily programmable
- Timings, temperature, cycles
- Tailor conditions to your sequence

PCR 1 - Primer Selection

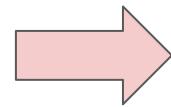


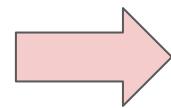


Amplicon

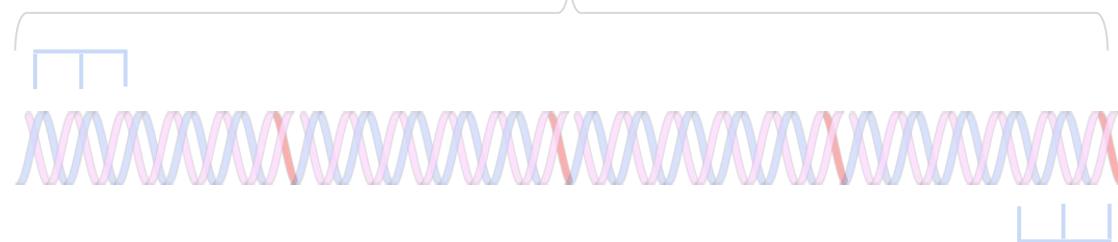


Amplicon length

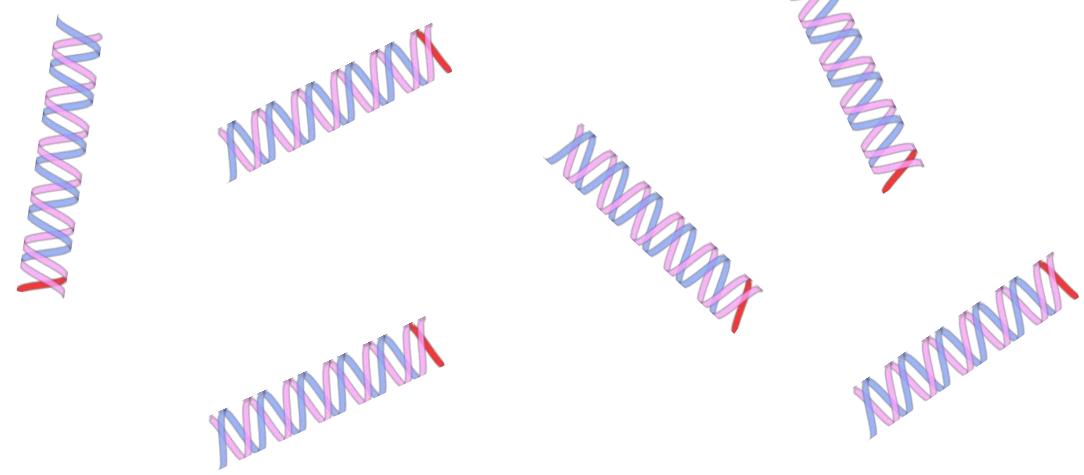
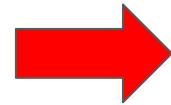




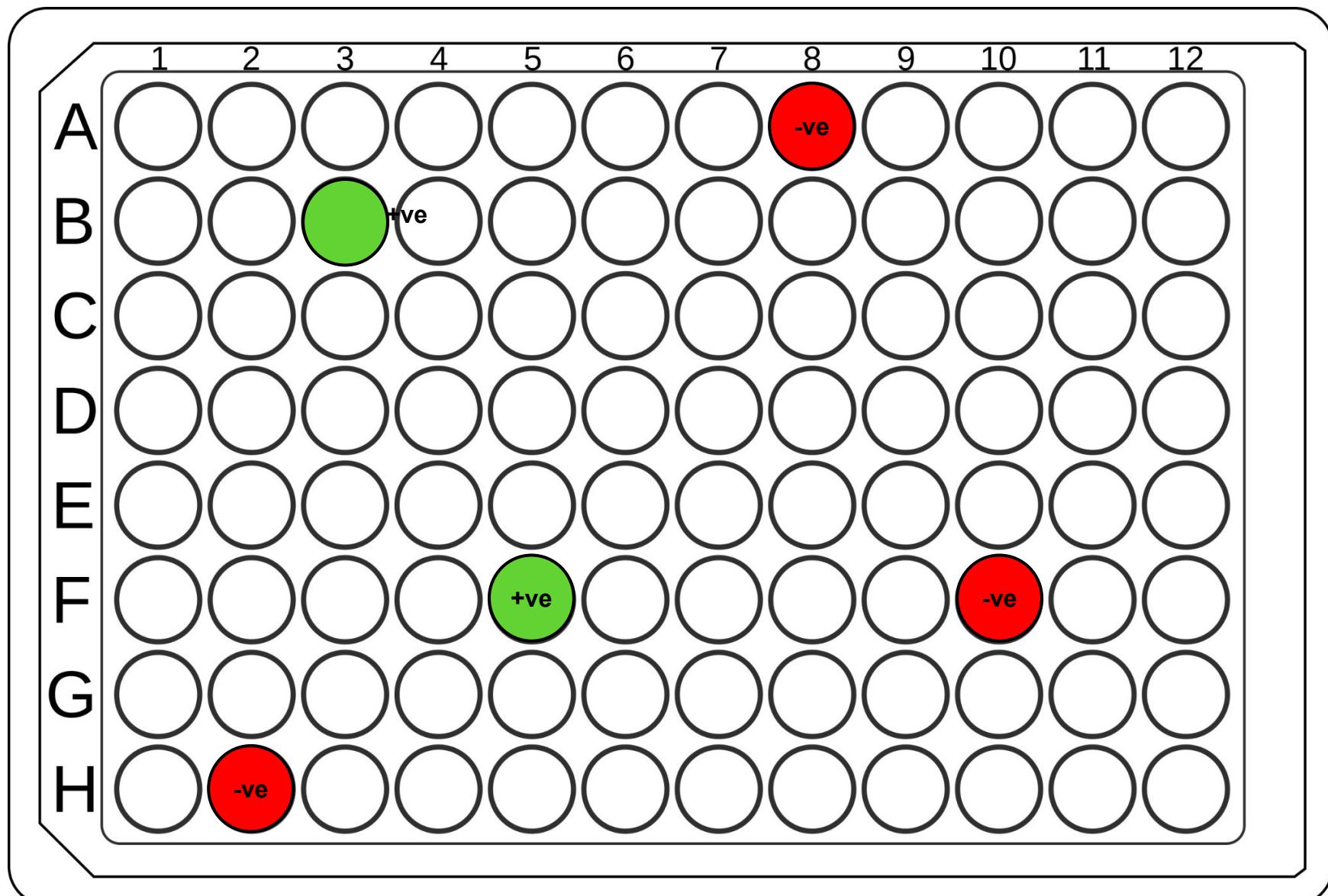
Amplicon



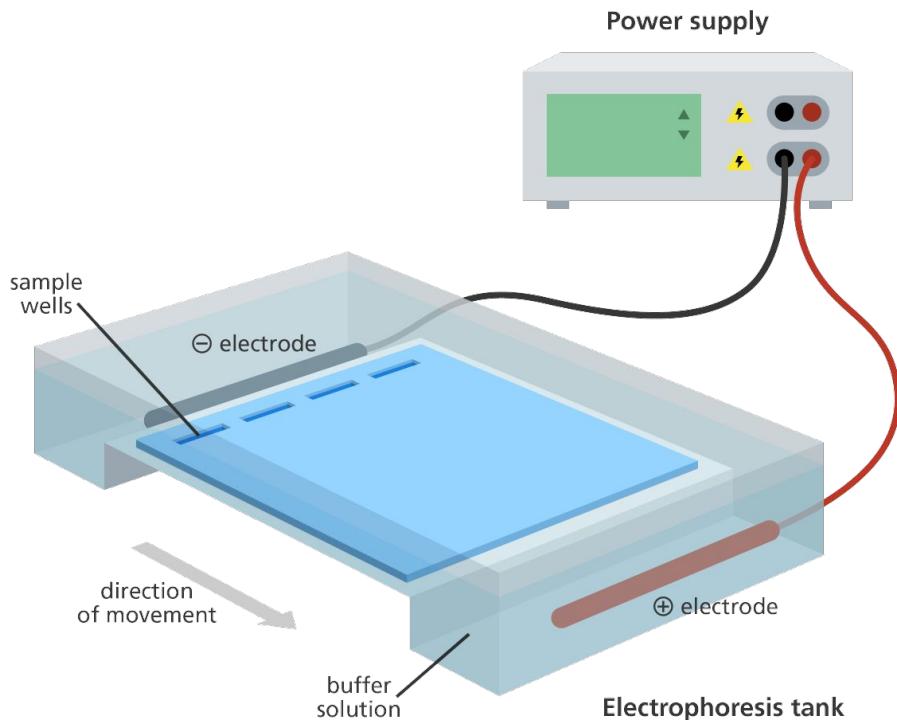
Amplicon length



PCR 1 - Quality control



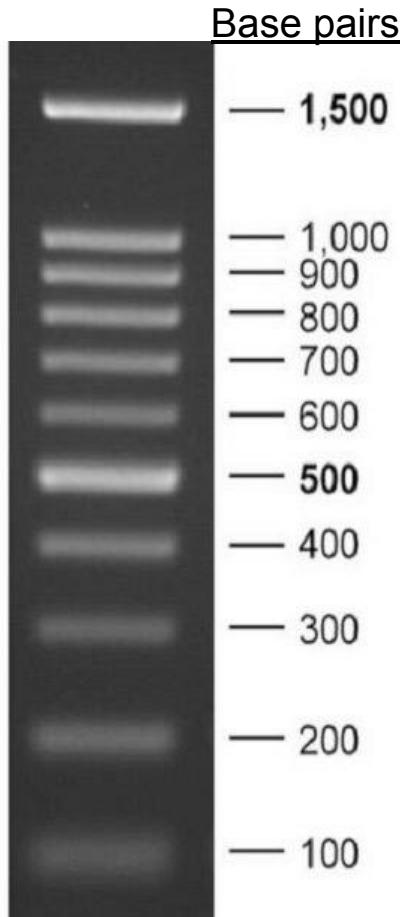
PCR 1 - Gel check



- Running your PCR products on an agarose gel is a useful way to visualise your products post-amplification
- The PCR products are added to the gel in lanes, and drawn through the gel by an electrical current

PCR 1 - Gel check

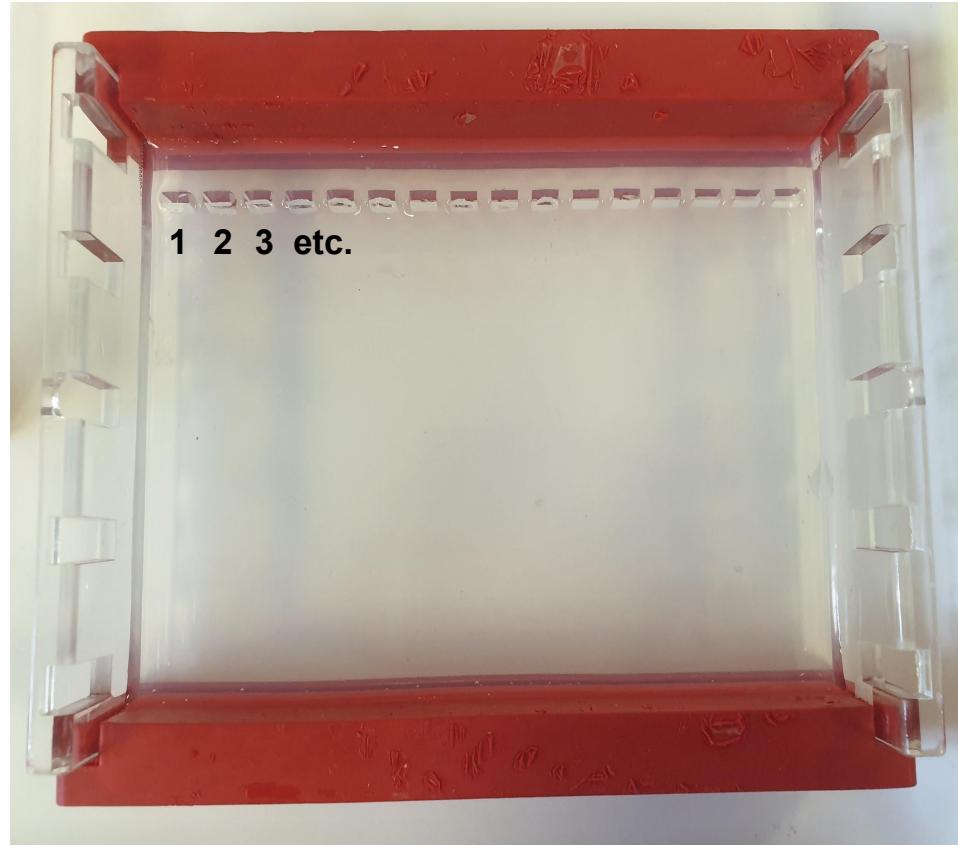
100bp ladder



- A ladder is usually added to the first lane in order to visualise the length of your PCR products, measured in base pairs
- We already have an idea of what the length our PCR products should be, based on the primers we designed

Time to load your gel

1. Remove red gel bumpers and carefully lower the tray and gel into the tank.
2. Add **5µl of the 100bp ladder** to the first well, being careful not to pierce the gel.
3. Pipette mix and add **10µl of the PCR product mix** to a well.
4. Repeat step 3 adding to the next well.
5. Pass to the next group member.

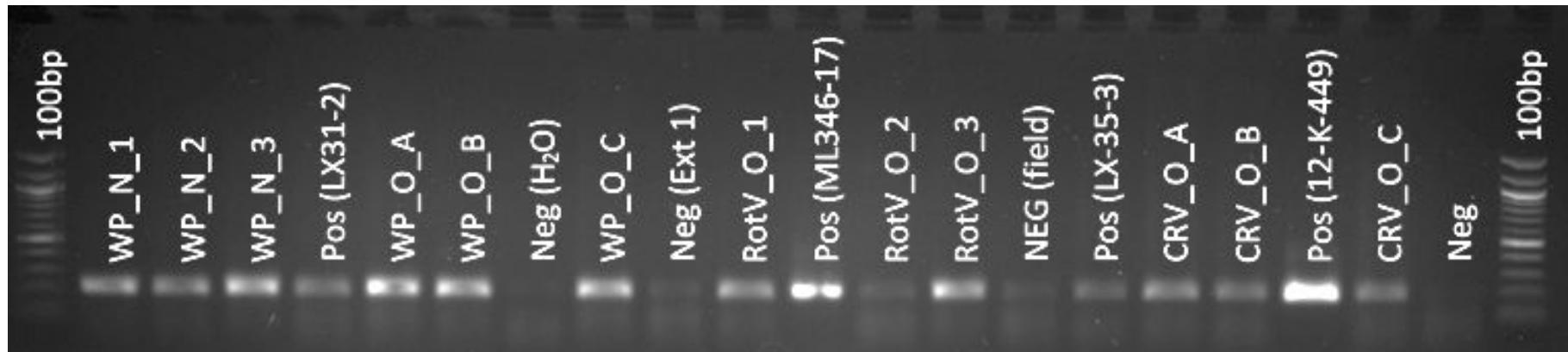


PCR 1 - Visualising products



A gel imaging dock like this uses UV light to illuminate the PCR products in the gel

PCR 1 - Visualising products

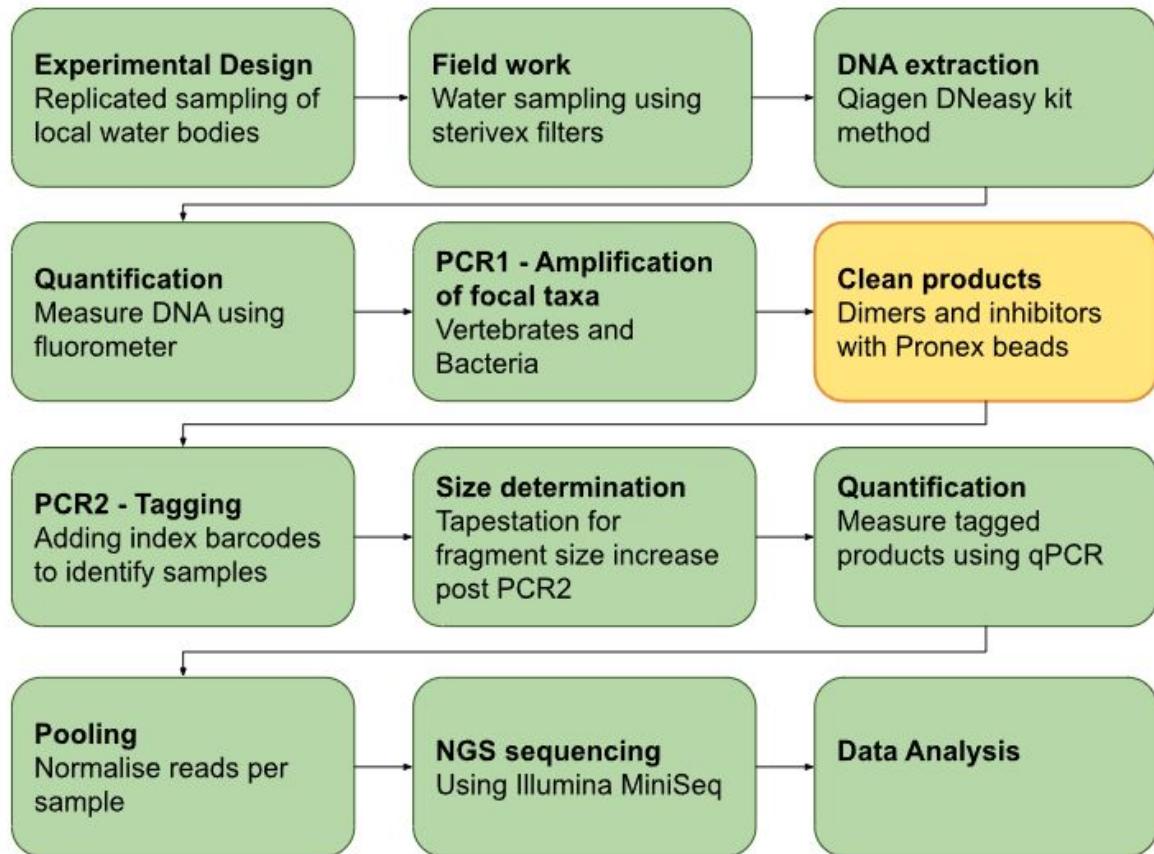


Post-PCR1 amplicons from a range of sampling sites.

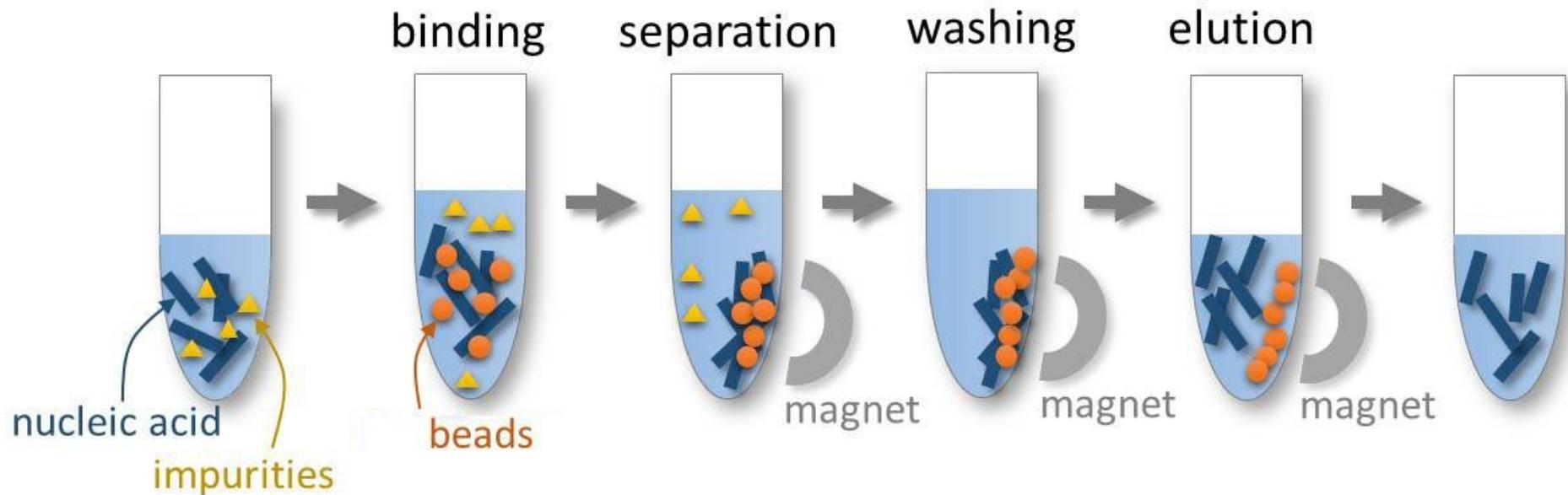
- The resulting image shows our PCR product as white bands
- We compare these bands to the ladders at either end
- This tells us the approximate size of our product
- The negative control should appear blank so long as no contamination has occurred

Clean & purify

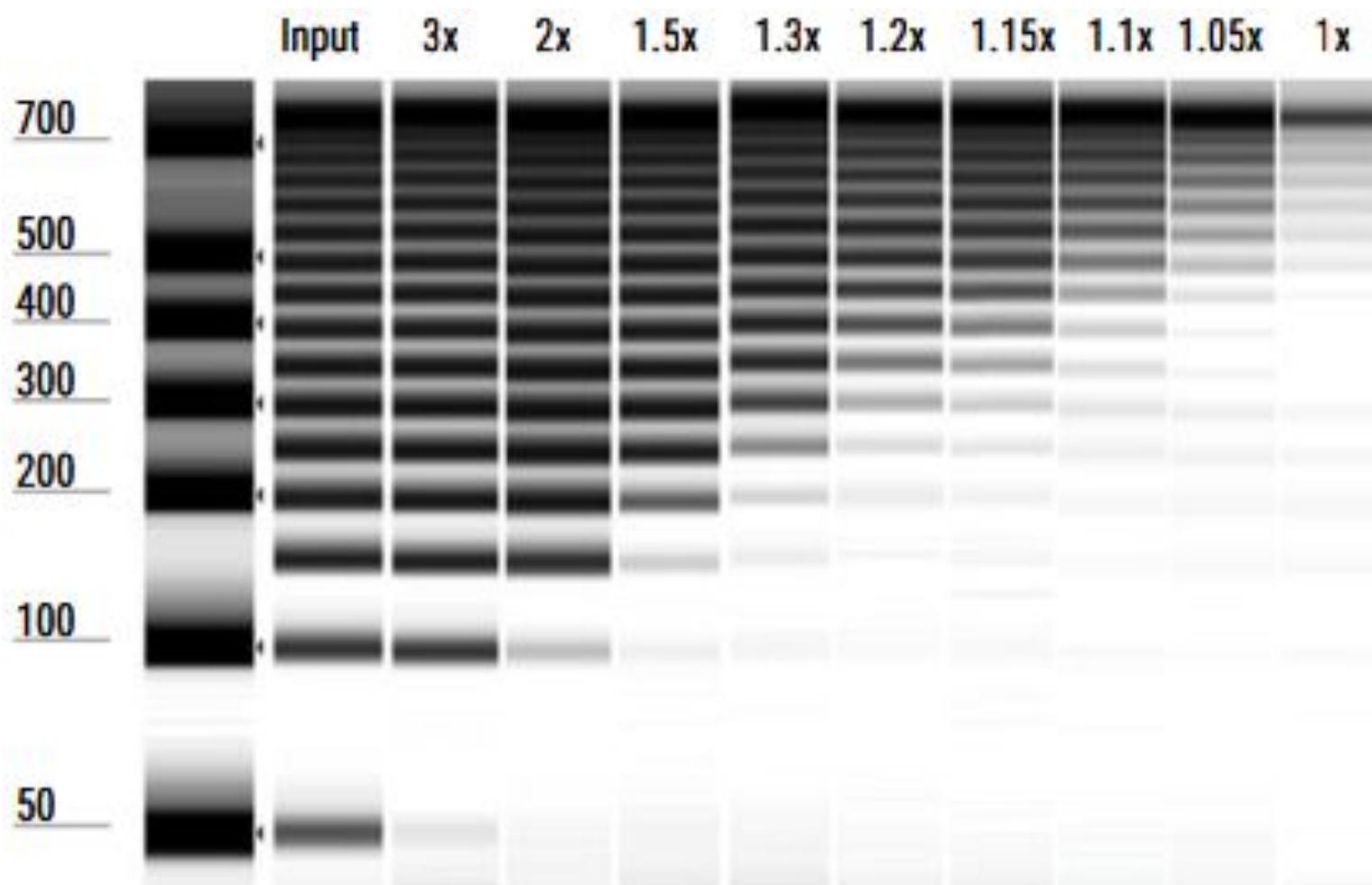
Clean PCR products of dimer and inhibitors using magnetic beads



PCR 1 - Bead cleaning

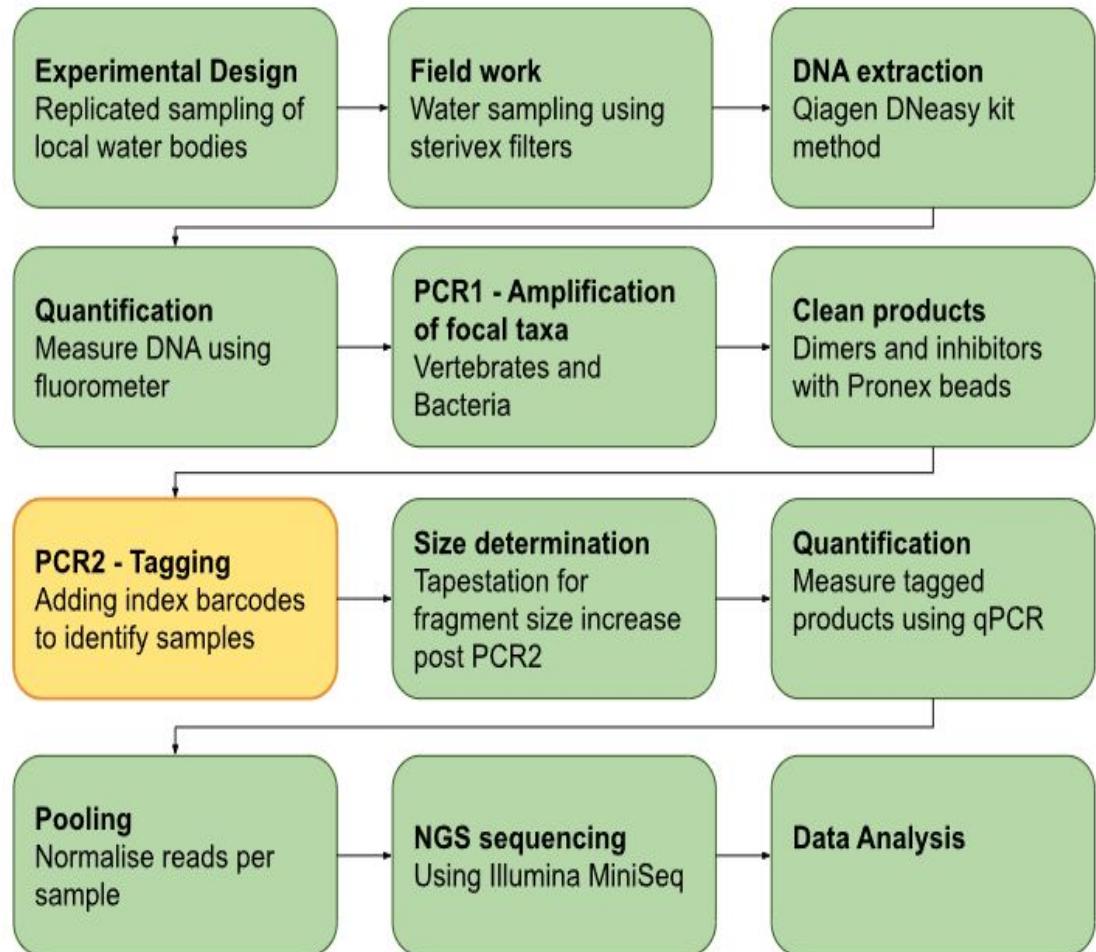


PCR 1 - Bead cleaning

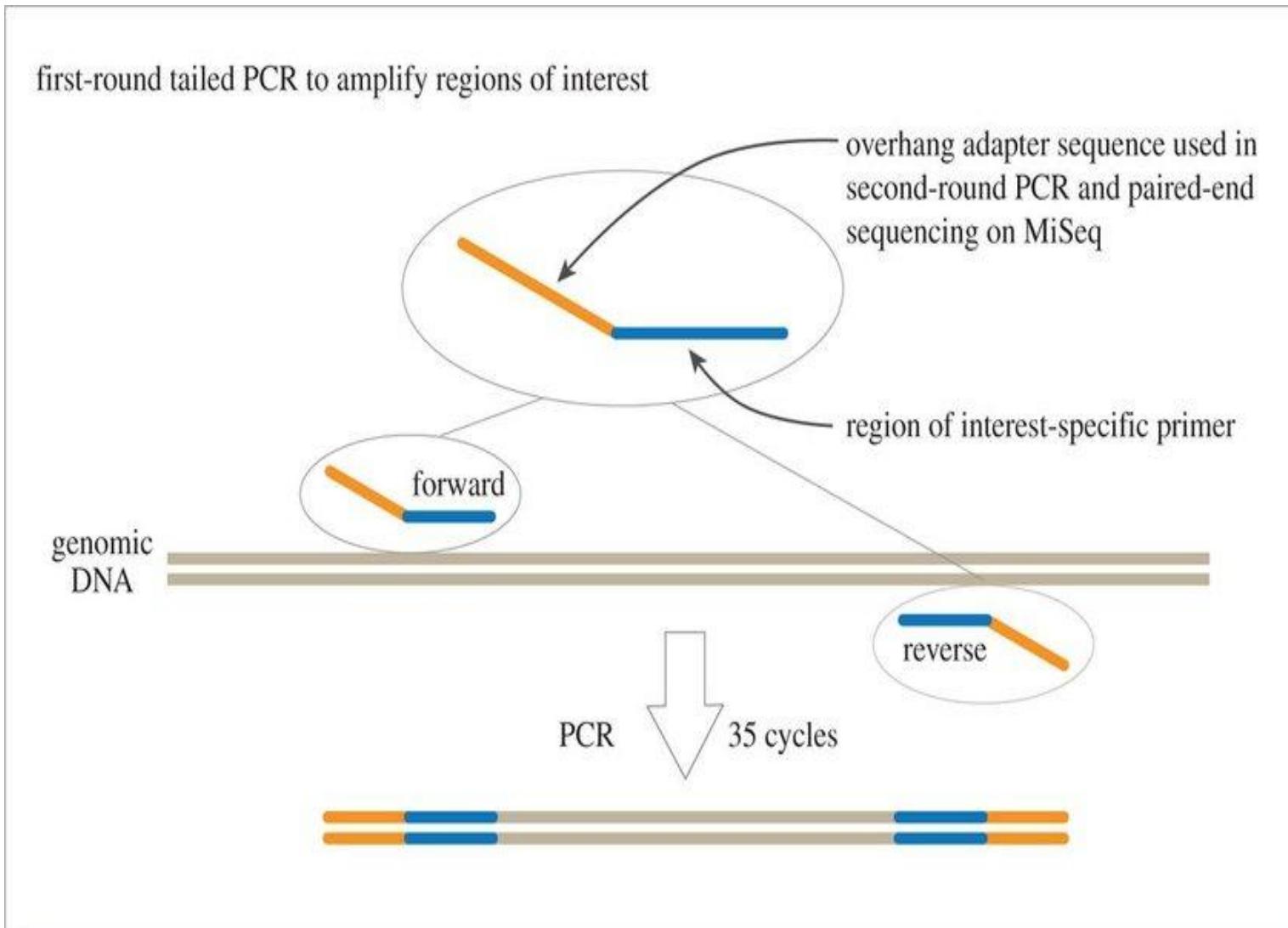


PCR 2

Tagging your amplicons with index sequences so samples can be identified once pooled for running on the sequencer

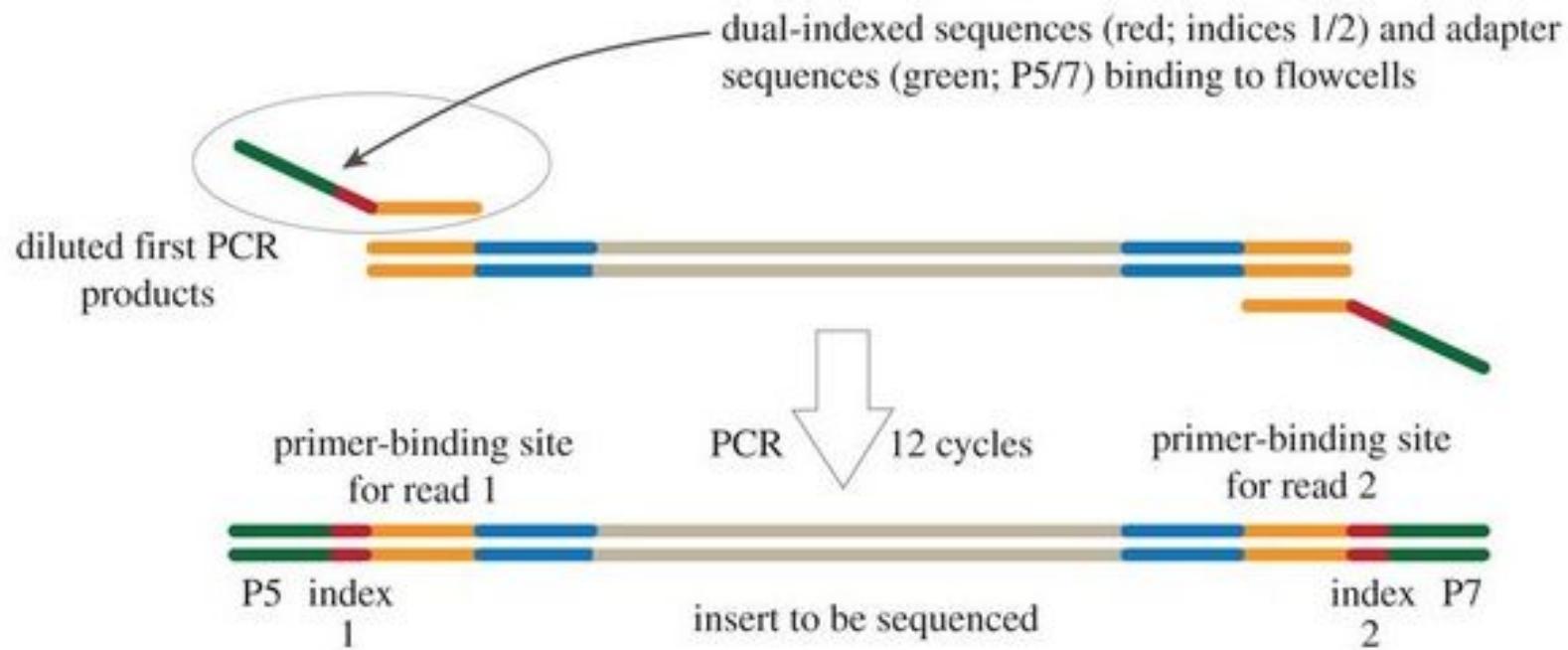


- During PCR1, our primers incorporate an overhang adapter sequence at the ends of our amplicons
- These adapters bind to unique identifiers, attached in a second PCR reaction, PCR2



PCR 2

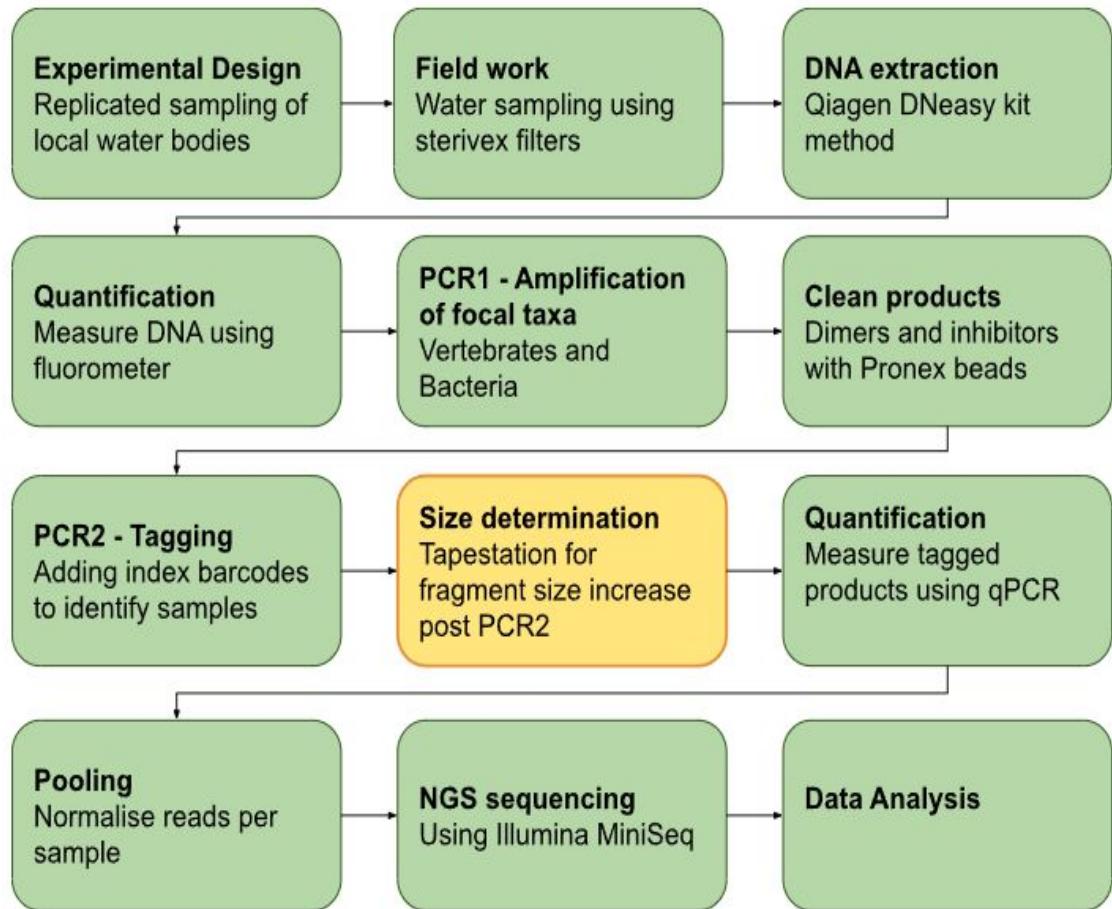
- In this second PCR reaction, unique identifier sequences are attached to our PCR 1 products



- These sequences allow individual samples to be identified in the bioinformatic data produced post-sequencing

Confirming fragment Size

Check that the PCR2 indexes have been correctly attached to the PCR product

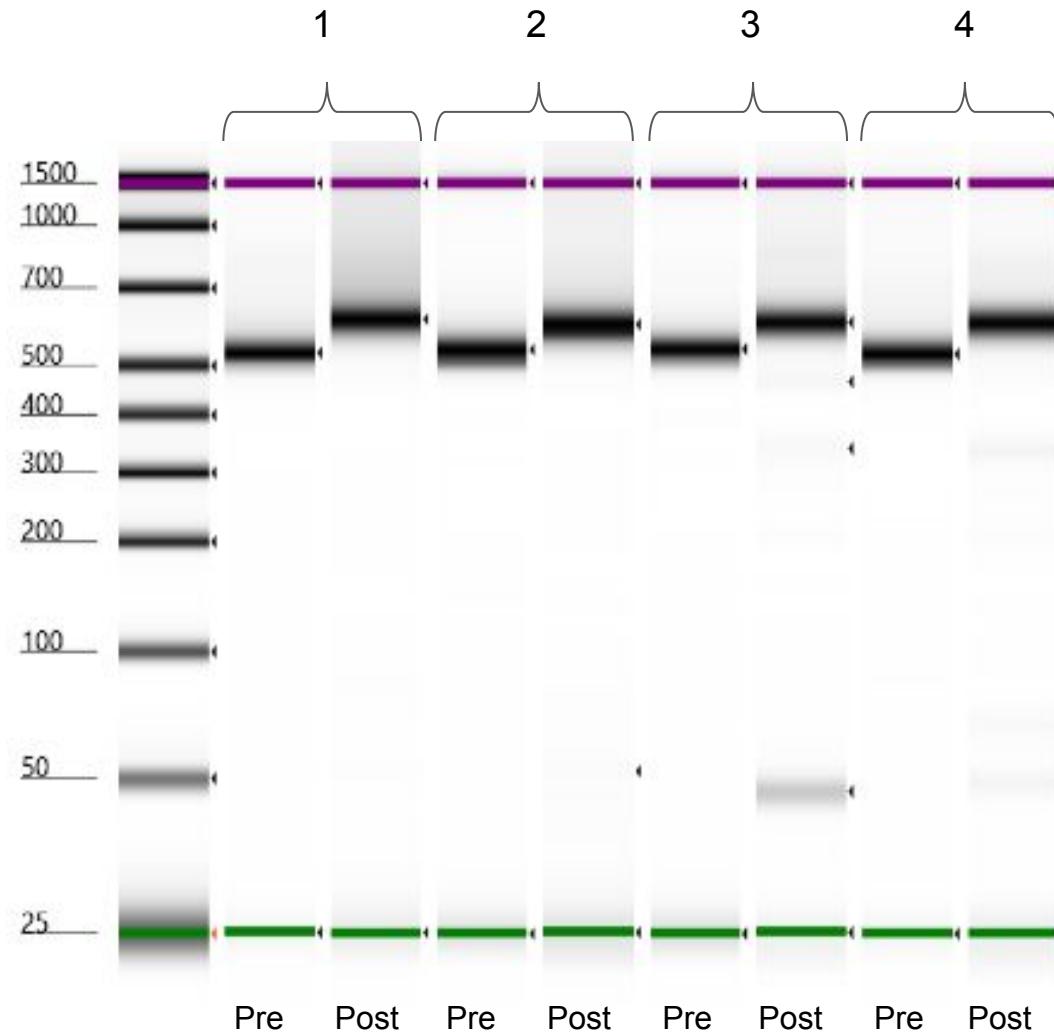


PCR 2 - Tapestation



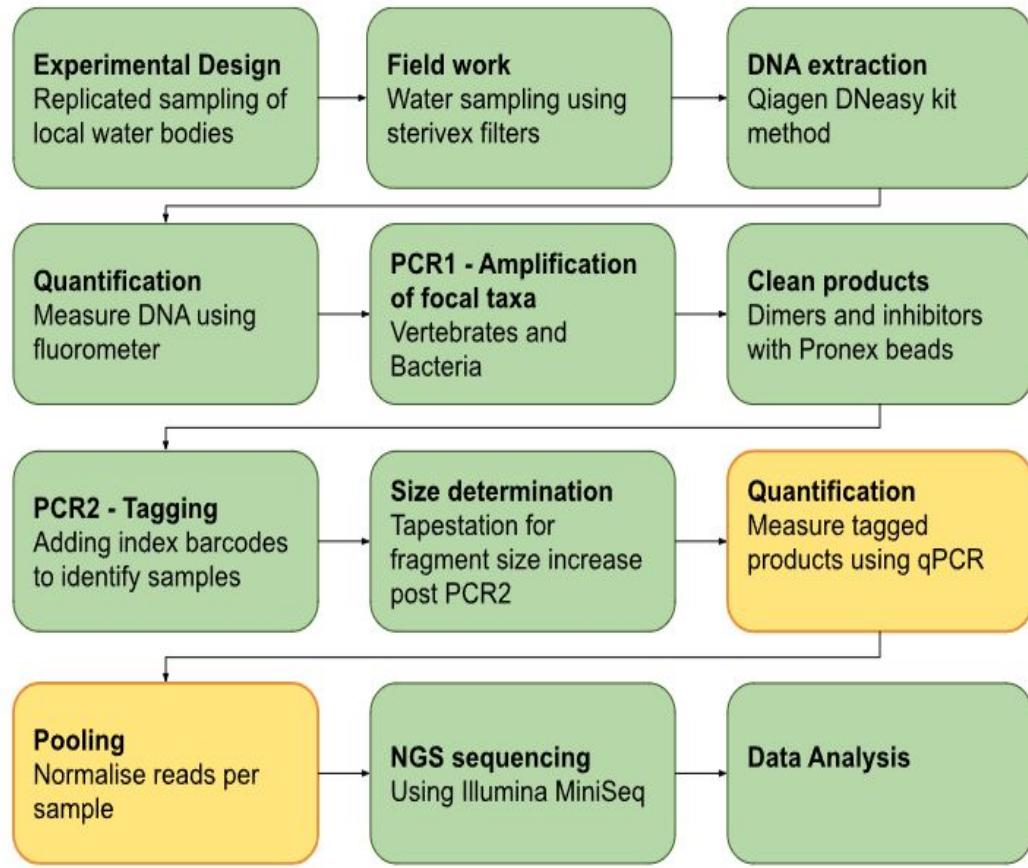
- Works like a miniature gel
- Compare samples pre and post-PCR 2
- Amplicons increase in size by length of unique identifier

PCR 2 - Tapestation

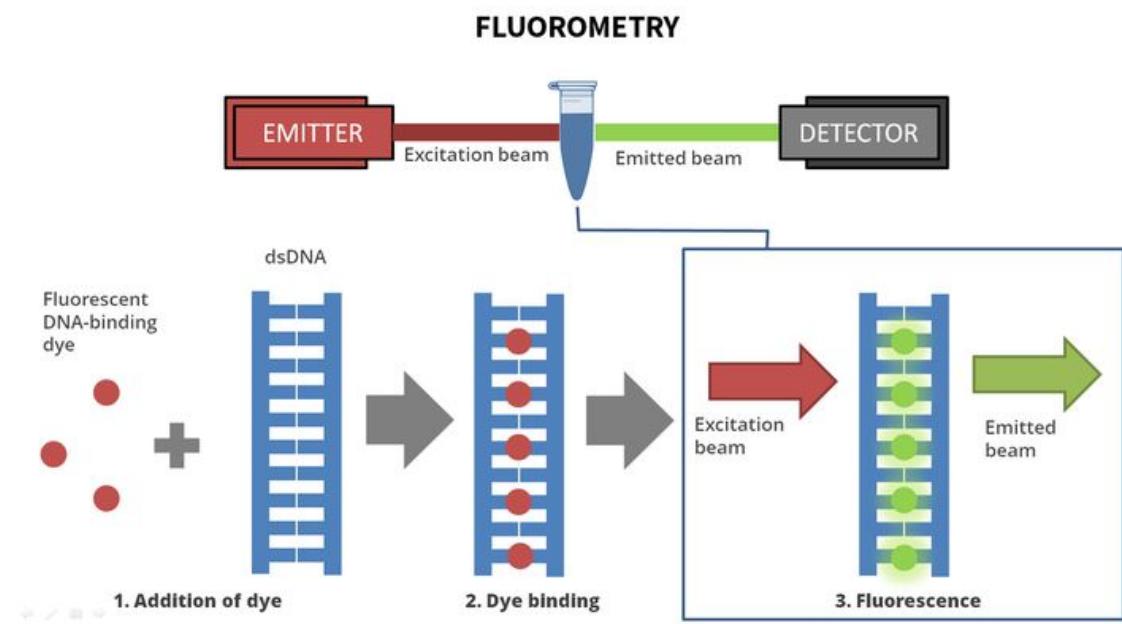


Quantify, pool,clean and quantify again.

Use fluorometry to
quantify products post
PCR, normalise and pool
for qPCR



PCR 2 - Quantification



Pooling: For an even representation of all samples in the final data

$$\frac{(\text{Final Volume } (\mu\text{l}) \times \text{Final Concentration } (\text{ng}/\mu\text{l}))}{(\text{Number of Samples} \times \text{Initial Concentration } (\text{ng}/\mu\text{l}))} = \text{Volume to Pool } (\mu\text{l})$$

$$X_{02} = 13.39$$

$$X_{03} = 6.87 \quad \text{Total Volume to Pool } (\mu\text{l}) = 42.87$$

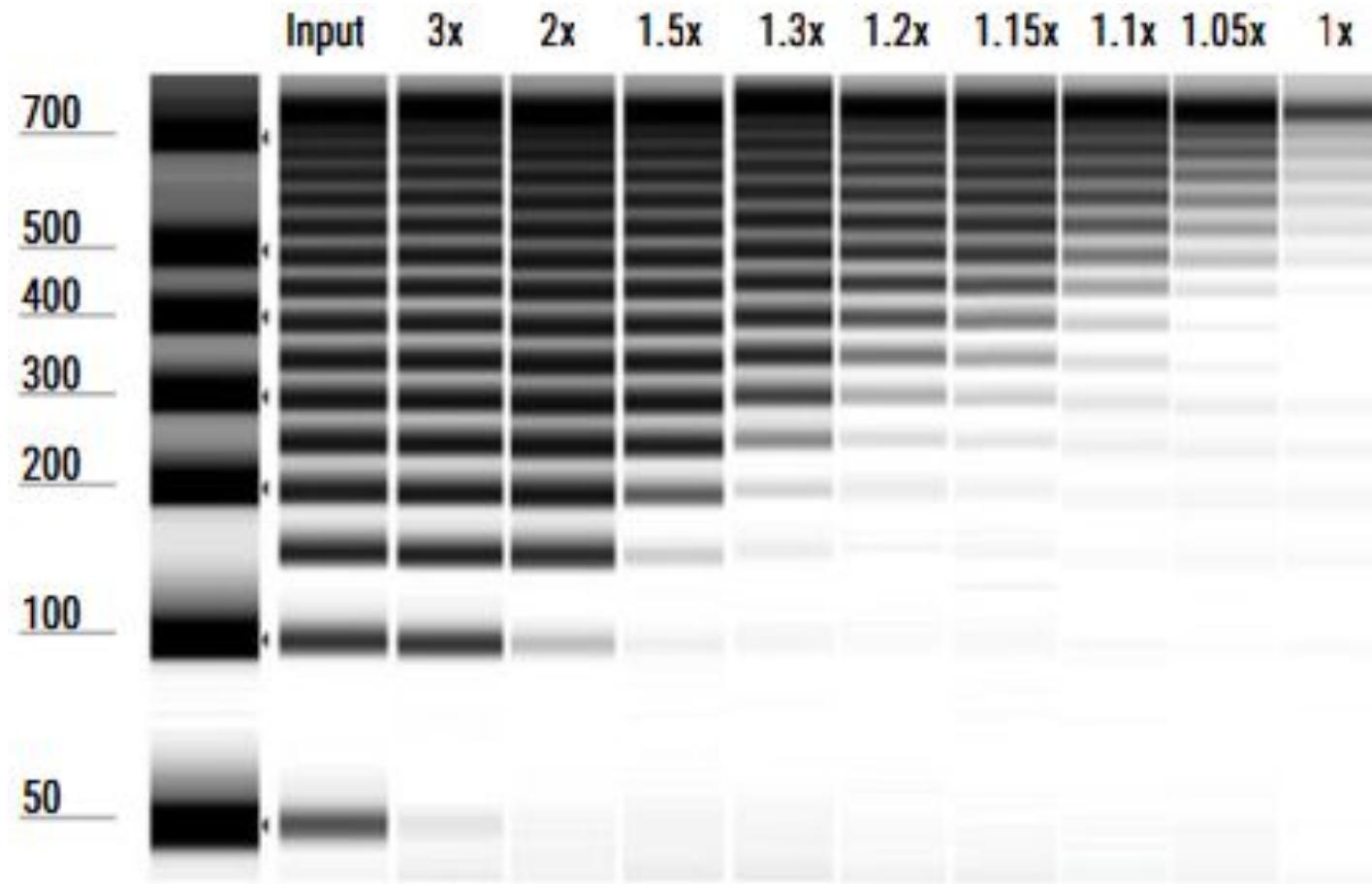
$$X_{04} = 8.31$$

$$\text{Volume of H}_2\text{O } (\mu\text{l}) = 7.13$$

$$X_{05} = 3.65$$

$$X_{06} = 5.81$$

PCR 2 - Bead cleaning

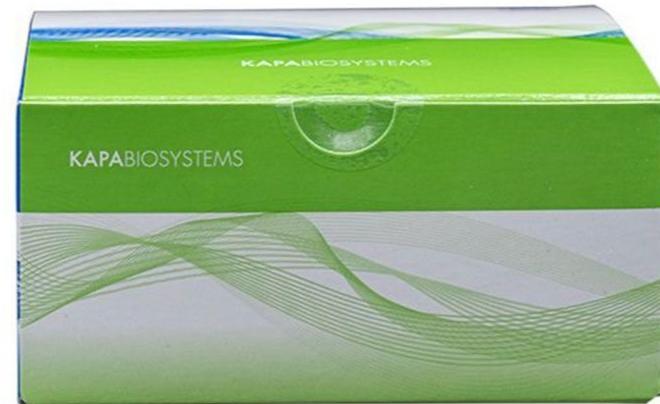


Quantification - qPCR



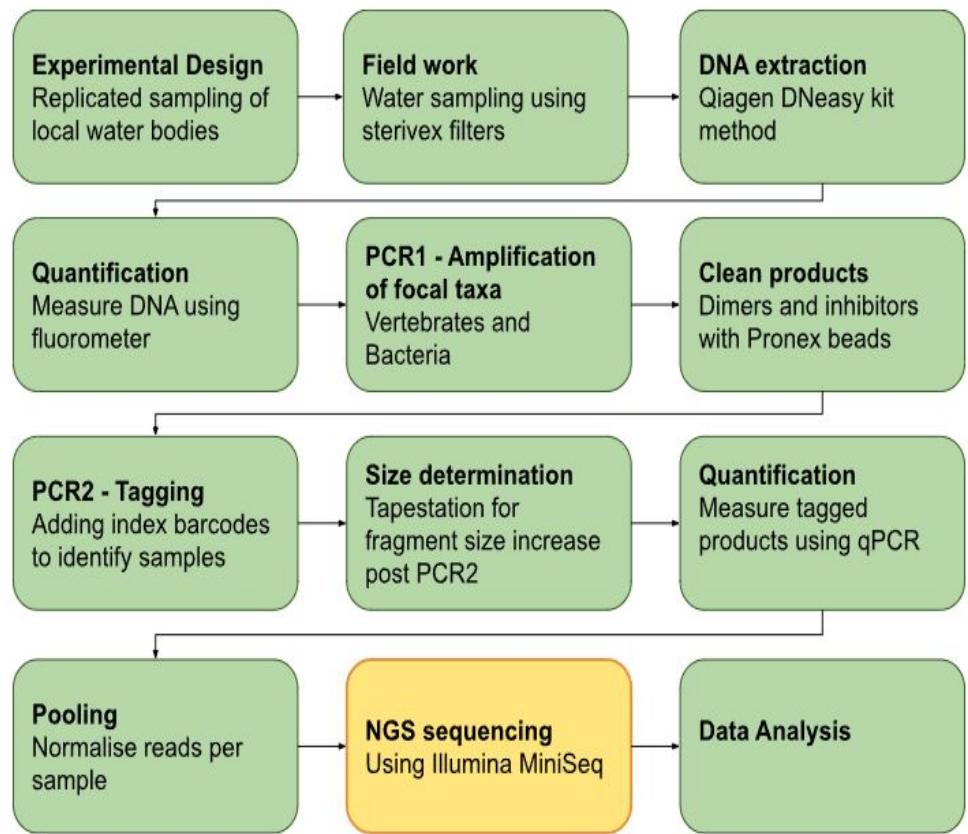
Due to the presence of unlabelled DNA from PCR2 we must quantify products by qPCR

Using a commercial kit we can quantify products for Illumina sequencing on the qPCR machine that only target DNA with i5 and i7 indexes (identifier sequences) attached



Sequencing the amplicon library

Running the
library on the
Sequencer (e.g.
Illumnia
MiniSeq)



Sequencing - Illumina NGS platforms



Different sequencing platforms can be used depending on the number of reads you require but also the length of the products you want to sequence

Different Illumina NGS platform specifications

Platform	Read length	Output
iSeq 100	2 x 150bp	1.2 Gb
MiniSeq	2 X 150bp	2.1 -2.4 Gb, 6.6-7.5 Gb
MiSeq	2 x 150bp, 2 x 250bp, 2 x 300bp	4.5-5.1 Gb, 7.5-8.5 Gb, 13.2-15 Gb*
NovaSeq	2 x 150bp	200 - 3000 Gb

* Nano and micro kits also available, used for testing (300Mb - 1.2Gb data)

Previous eDNA study

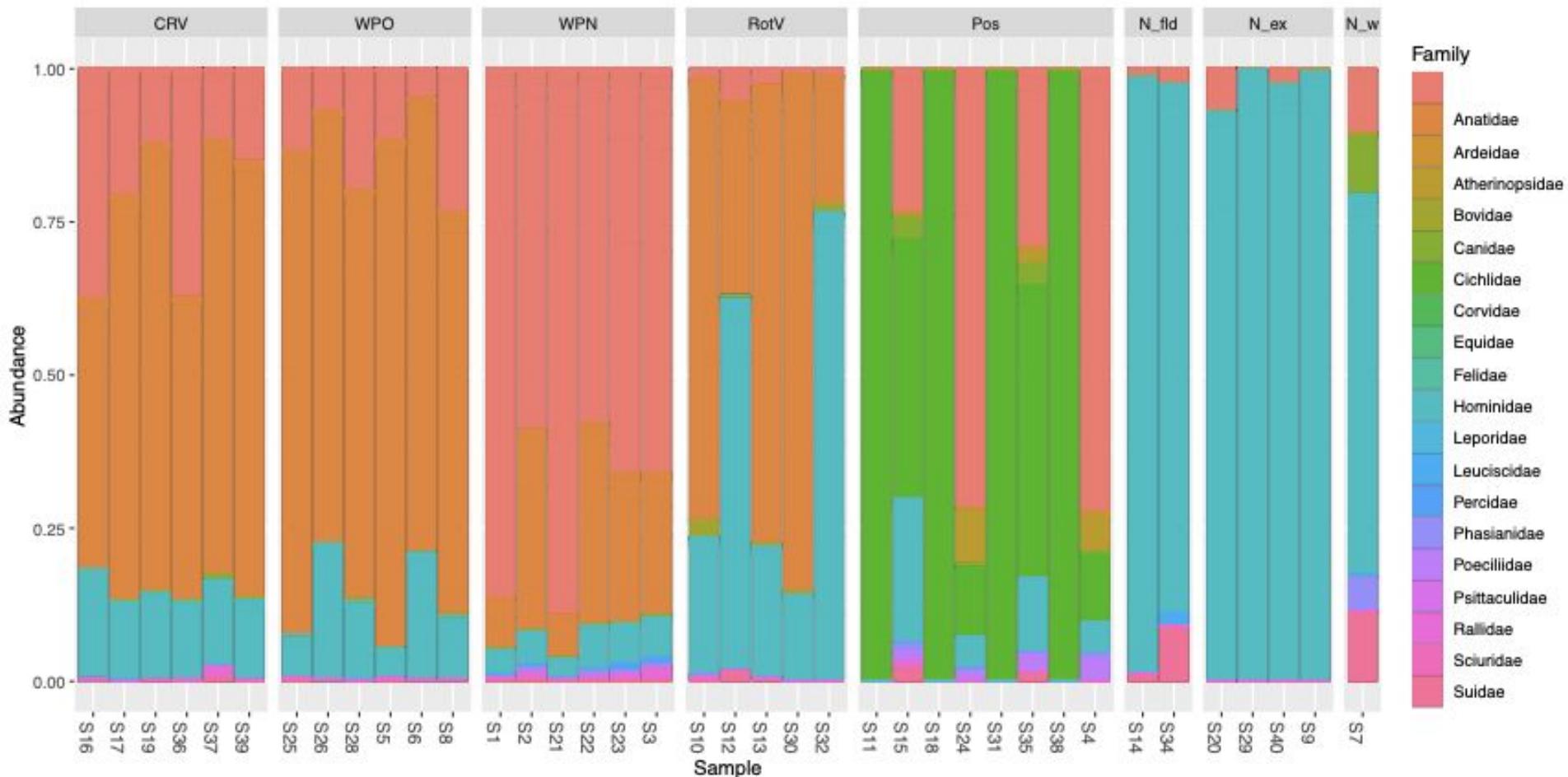
- Samples:
 - Weston Park November (**WP_N**) samples
 - Weston Park October (**WP_O**)
 - Rother Valley (**RotV_O**)
 - Crookes Valley (**CRV_O**)
 - Positive controls - Mexican river samples, African Cichlid stomachs (**Pos**)
 - Negatives
 - Field (**N_fld**)
 - Extraction (**N_ex**)
 - PCR (**N_w**)
- Riaz 12S-V5 primer set
(<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3241669/>)



What did we find?



What did we find?



Take home message...

Consider this equation when planning your metabarcoding projects

$$\frac{\text{Average Yield from the Sequence Run}}{\text{Read Depth} \times \text{Number of Primers}} = \text{Number of Reactions}$$