

16SrRNA Intermediate Bioinformatics Online Course: Int\_BT\_2019

## Module 3:

# Sample collection, extraction and library preparation for 16S NGS analyses







## **Learning Outcomes**



Describing different aspects of planning for 16S experiments; for example study design, DNA extraction methods and laboratory workflows.

- Understand what is meant by the term "16S rRNA gene" and why we are interested in this gene.
- Understand the concepts behind Sanger sequencing and highthroughput sequencing
- Know why it is important to clearly plan an experiment and how different components to the study and experimental design may influence data generated and downstream analyses.
- Know that techniques other than 16S rRNA sequencing are also available to study microbial profiles.







16SrRNA Intermediate Bioinformatics Online Course: Int\_BT\_2019

## Module 3:

## Sample collection, extraction and library prep for 16S NGS analyses

**Part 3.1** 

16S rRNA high throughput sequencing: how it works







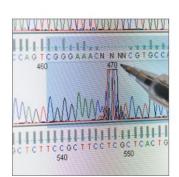




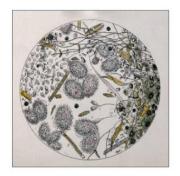
mid-1600s: First microbes described



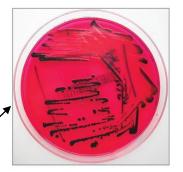
1800s - Present: Culture, staining, and microscopy used to study microbes that can be cultured



circa 1600: Microscope invented



1800s: Connection made between microbes and disease



1990s:
DNA sequencing
becomes available,
allowing study of
microbes that cannot
be cultured

Advantages: Study viable organisms

**Disadvantages**: Thought that less than 1%

of all bacterial species are cultivatable









#### **Culture-independent molecular methods:**

Marker gene-dependent

Fingerprinting techniques

Sanger sequencing

High-throughput sequencing

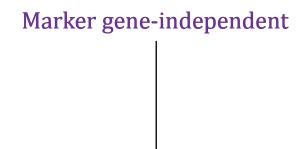
16S rRNA gene

All organisms need ribosomes to make protein

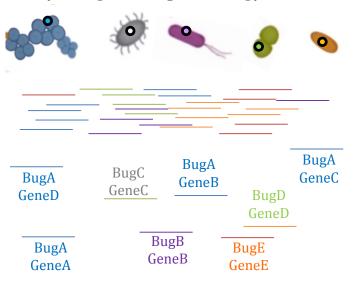
Any gene that makes up ribosome may be a good marker gene

Ribosomal RNA (rRNA) never gets transcribed to protein

Gene focussed on: **16SrRNA gene** (~1500bp)



High-throughput sequencing (shotgun sequencing)



https://www.youtube.com/watch?v=8Aa\_mnyXm70

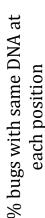
16SrRNA Intermediate Bioinformatics Online Course:
Int\_BT\_2019
Shantelle Claassen-Weitz

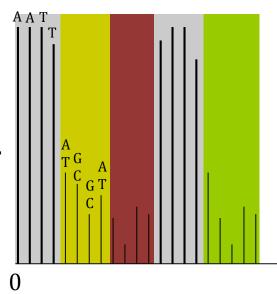












Some parts of the gene are structurally very important for the ribosome to work =  $\underline{\text{conserved regions}}$ 

Some parts of gene have *mutations* but the ribosome still works = <u>variable regions</u>

~1500 nucleotides



The entire gene is too long to sequence without errors for most DNA sequencers

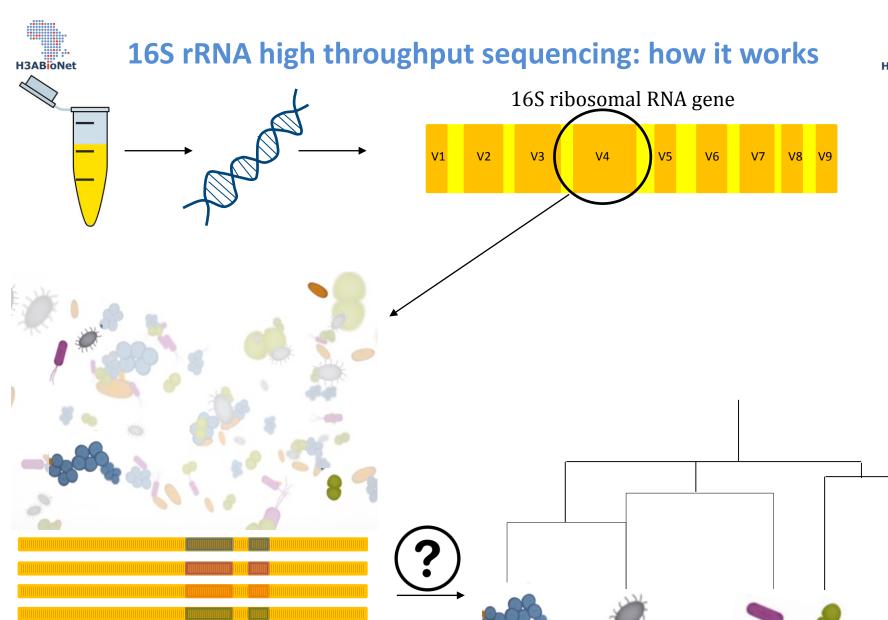
Find a portions with enough variability to distinguish species

https://www.youtube.com/watch?v=8Aa\_mnyXm70





16SrRNA Intermediate Bioinformatics Online Course:
Int\_BT\_2019
Shantelle Claassen-Weitz



DNA amplicons

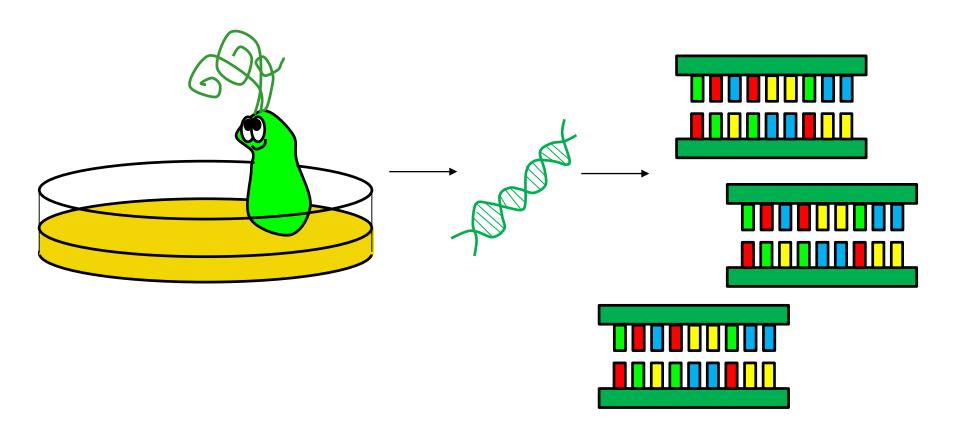
Pan African Bioinformatics Network for H3Africa

**H3ABioNet** 







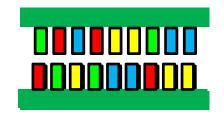










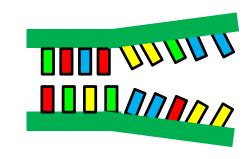










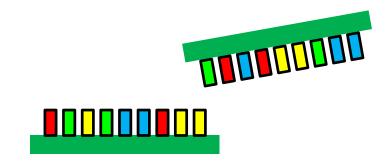










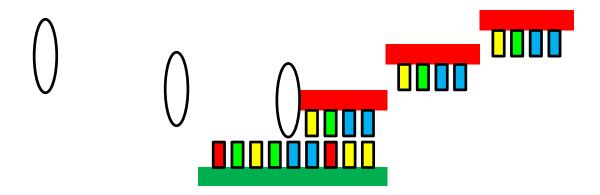










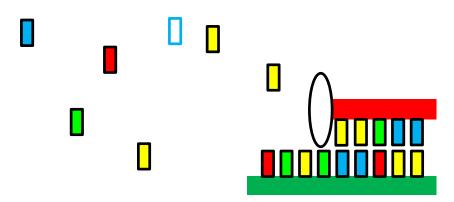










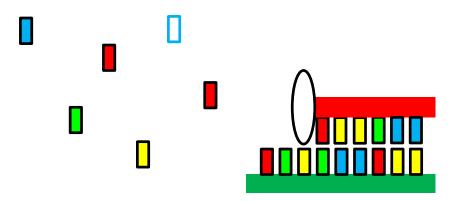










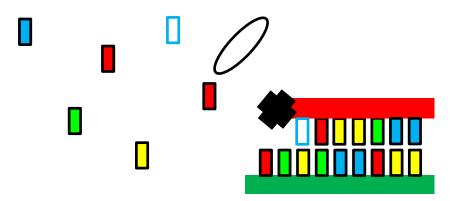










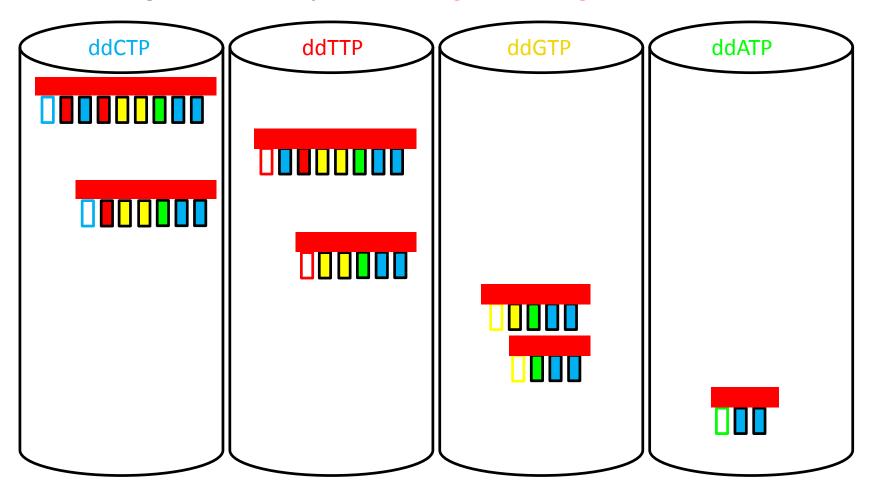










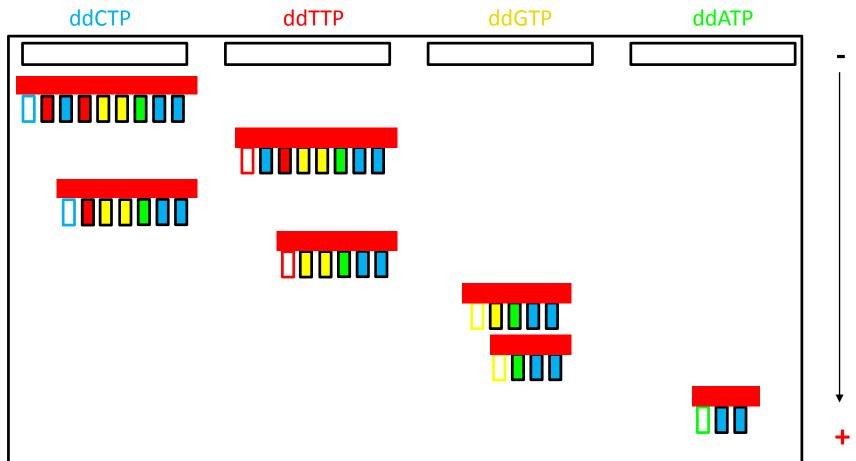










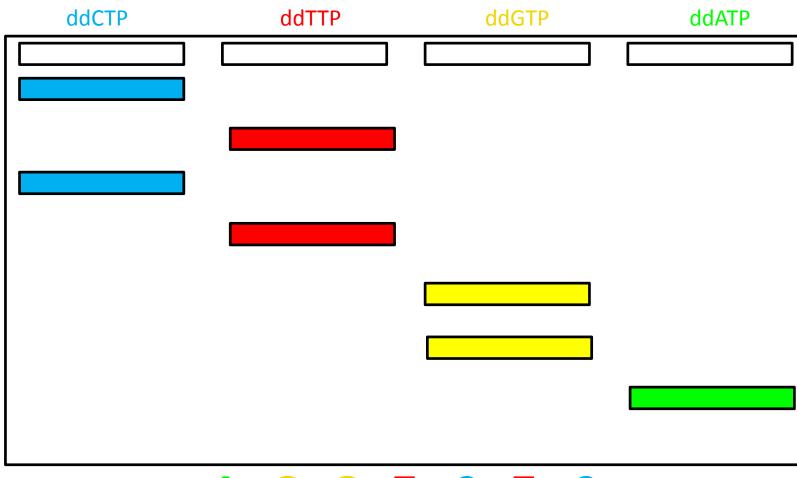
















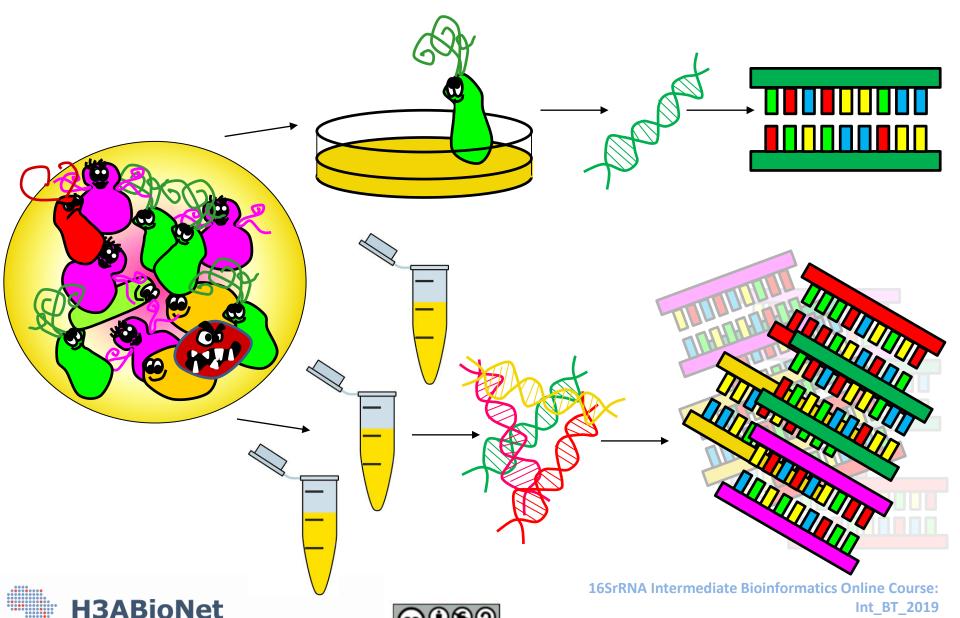


**Pan African Bioinformatics Network for H3Africa** 

## 16S rRNA high throughput sequencing: how it works



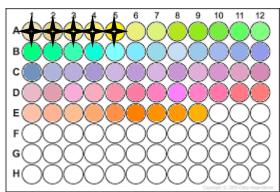
**Shantelle Claassen-Weitz** 





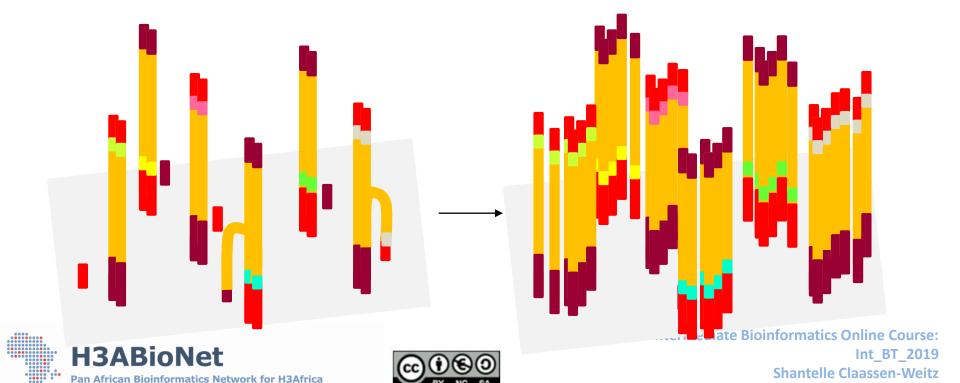


#### DNA from a number of samples



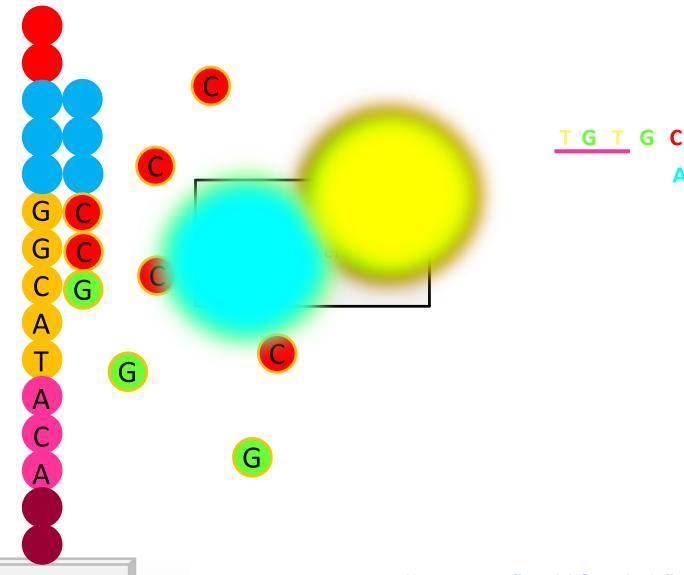
Undergo amplification using a multiplex approach



















#### In summary:

- 16S rRNA gene is by far the most common housekeeping gene targeted to study bacterial phylogeny and classification.
- High throughput sequencing targeting the 16S rRNA gene allows for large quantities of DNA to be sequenced much more quickly and cheaply compared to Sanger sequencing.
- High throughput sequencing of the 16S rRNA gene allows for identification as well as relative abundance quantification of all bacteria present in a sample.
- 16S rRNA high-throughput sequencing allows for processing multiple samples together via the use of barcoded (indexed) primers.



