Next Generation Sequencing Applications in Food Safety and Quality

Our science

- National and international centre of excellence for interdisciplinary investigation and problem solving across plant and bee health, crop protection, sustainable agriculture, food and feed quality and chemical safety in the environment.

- Remote, faster, early-warning diagnostics and interventions
- Discovery technology across the agri-food chain
- Smarter surveillance data and risk analysis
- Sustainability in the agri-food chain, increasing output and reducing waste

Facilities

500 staff including **350** scientists

80 acre secure site with field plots, glasshouses and a specialist wildlife unit

4km of lab benching

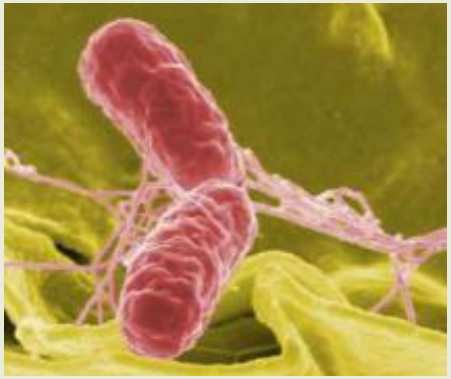
90 controlled environment facilities

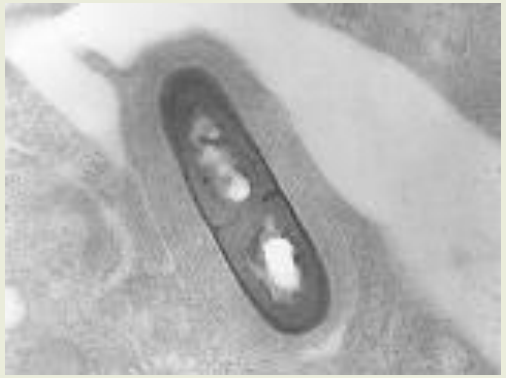
Containment – Cat 3 lab

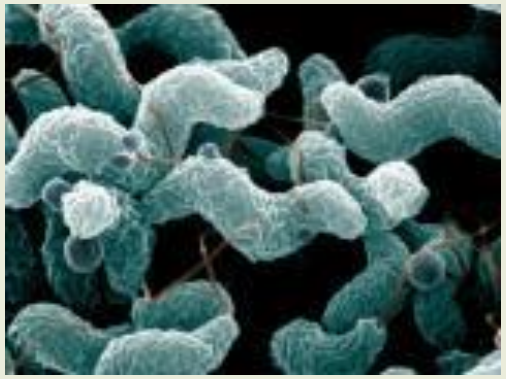
Latest chemical & molecular biological analysis tools

Foodborne pathogens

- 280,000 cases of *Campylobacter*, 30,000 cases of *Salmonella* in UK p.a.
- *Listeria monocytogenes* another FSA priority pathogen – case fatality rates 20 to 40%
- Food industry has additional concerns – brand, reputation, litigation etc.





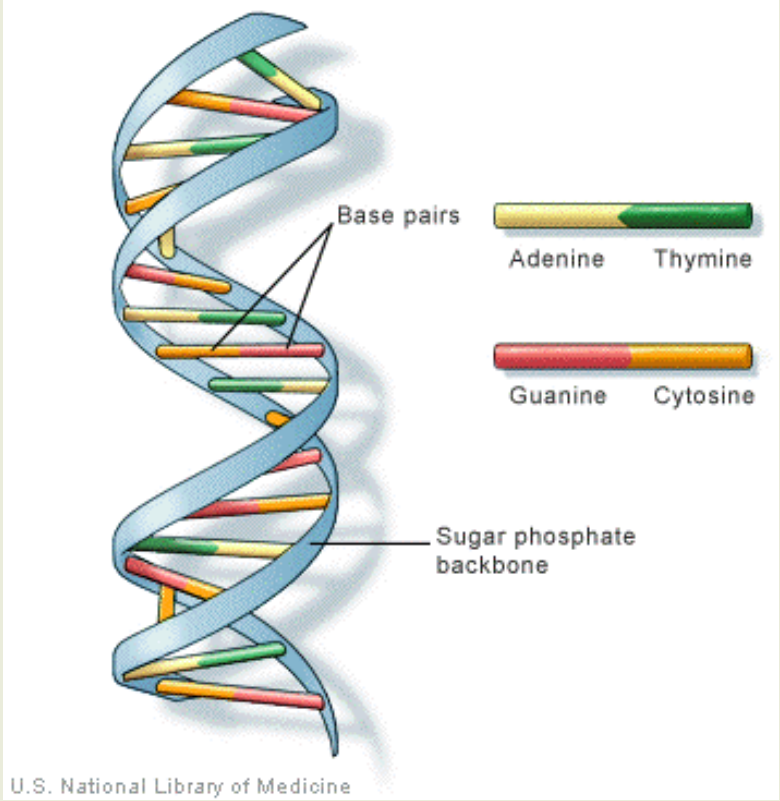


Subtyping bacteria

- By identifying bacterial species, and distinguishing subtypes within a species, we can link cases together and start to understand pathogen spread

DNA sequencing

- Genetic material in organisms encoded in a chemical polymer
- DNA sequencing reveals the order of base pairs along the DNA strand

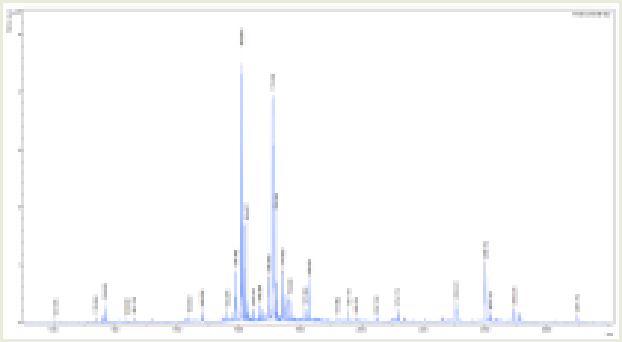


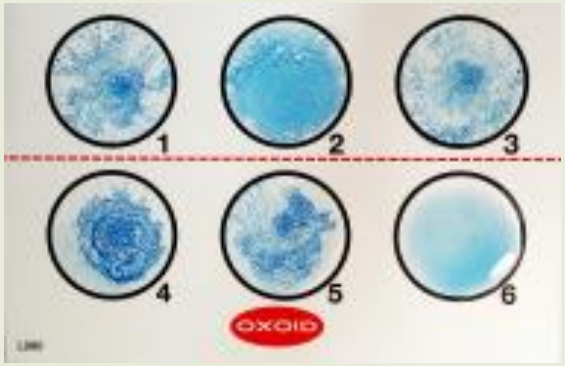
Why does the sequence matter?

- Genes encode proteins, the building blocks of cells
- Relatedness between organisms can be inferred from sequence similarity

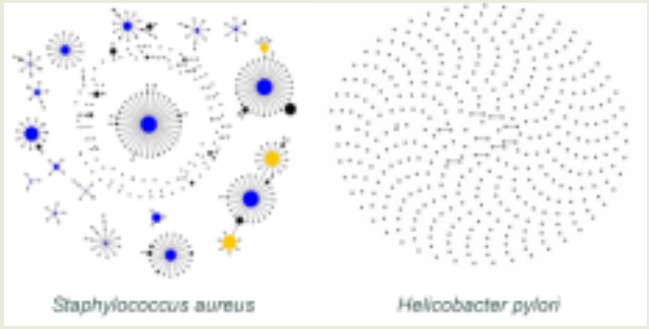
Ways to ID bacteria

- MALDI-TOF – rapid method for identifying species

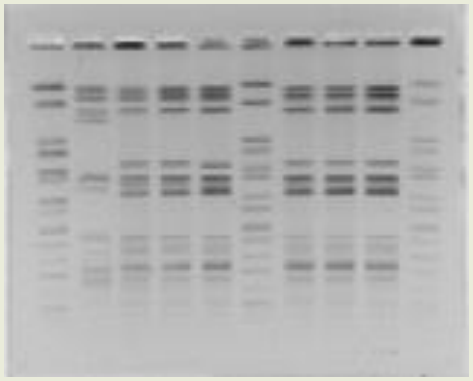




- Serotyping – identifies subtypes based on surface antigens



- MLST (Multi Locus Sequence Typing) – more discriminatory



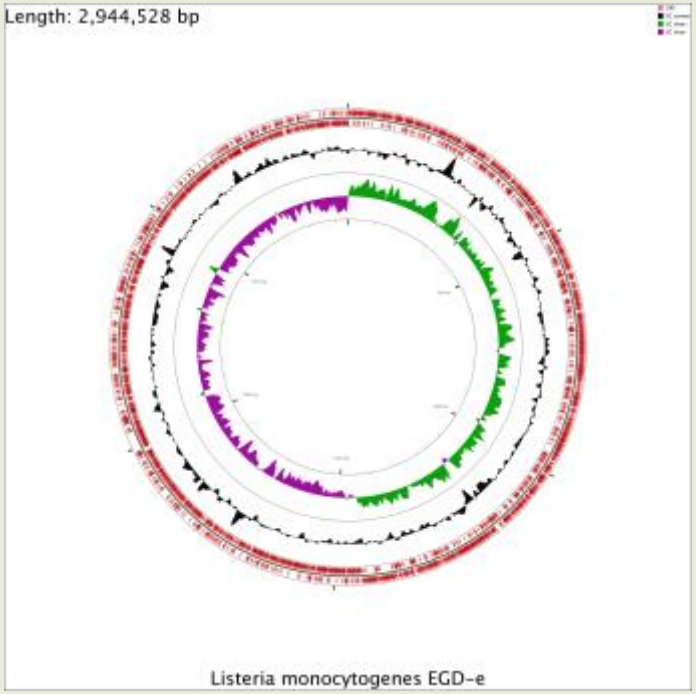
- PFGE (Pulsed Field Gel Electrophoresis) – highly discriminatory, but labour intensive

Whole Genome Sequencing (WGS)

- Conceptually very simple
- Extract total DNA from bacterial culture, sequence this in a massively parallel manner, interpret bioinformatically
- Bacterial genome sizes vary:
  - *Campylobacter jejuni* = 1,600,000 base pairs (1.6 Mb)
  - *Listeria monocytogenes* = 2.9 Mb
  - *Salmonella* Enteritidis = 4.7 Mb

- Coverage – to sequence

  L. monocytogenes at 20x coverage (to correct for sequencing errors) would require generation of 58,000,000 bases of DNA sequence.



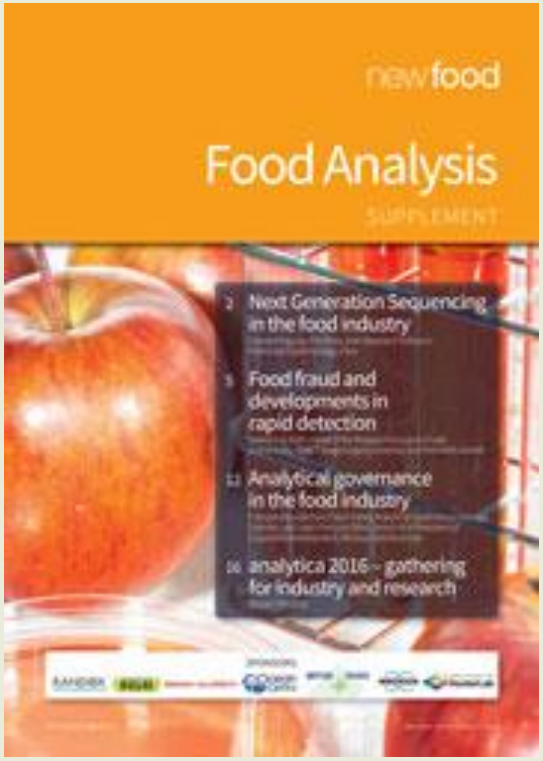
DNA sequencing technologies

Capillary based DNA sequencing (Life Tech) (<2007?)	GS-FLX-Titanium + (454) (2008-2016)	MiSeq (Illumina) (2013 -)	Oxford Nanopore MinION (2014 -)
96x750bp	750bp x 1 million	300bp x 2 x 25 million	4-50 kb+ x thousands
			
Long reads, Individual clones/barcodes	First out, Long reads, time consuming, expensive, phased out 2016	Simple work flow, relatively cheap, shorter reads, high quality data (worse at end)	Small, cheap, very long reads, improving error rate

Benefits to industry

- Identify multiple types vs continuous contamination with a resident strain
- Monitor trends over time
- Potential to implicate processes or equipment
- Reduce risk of recalls/demonstrate due diligence
- Suggest external sources of contamination – could require sharing anonymised data

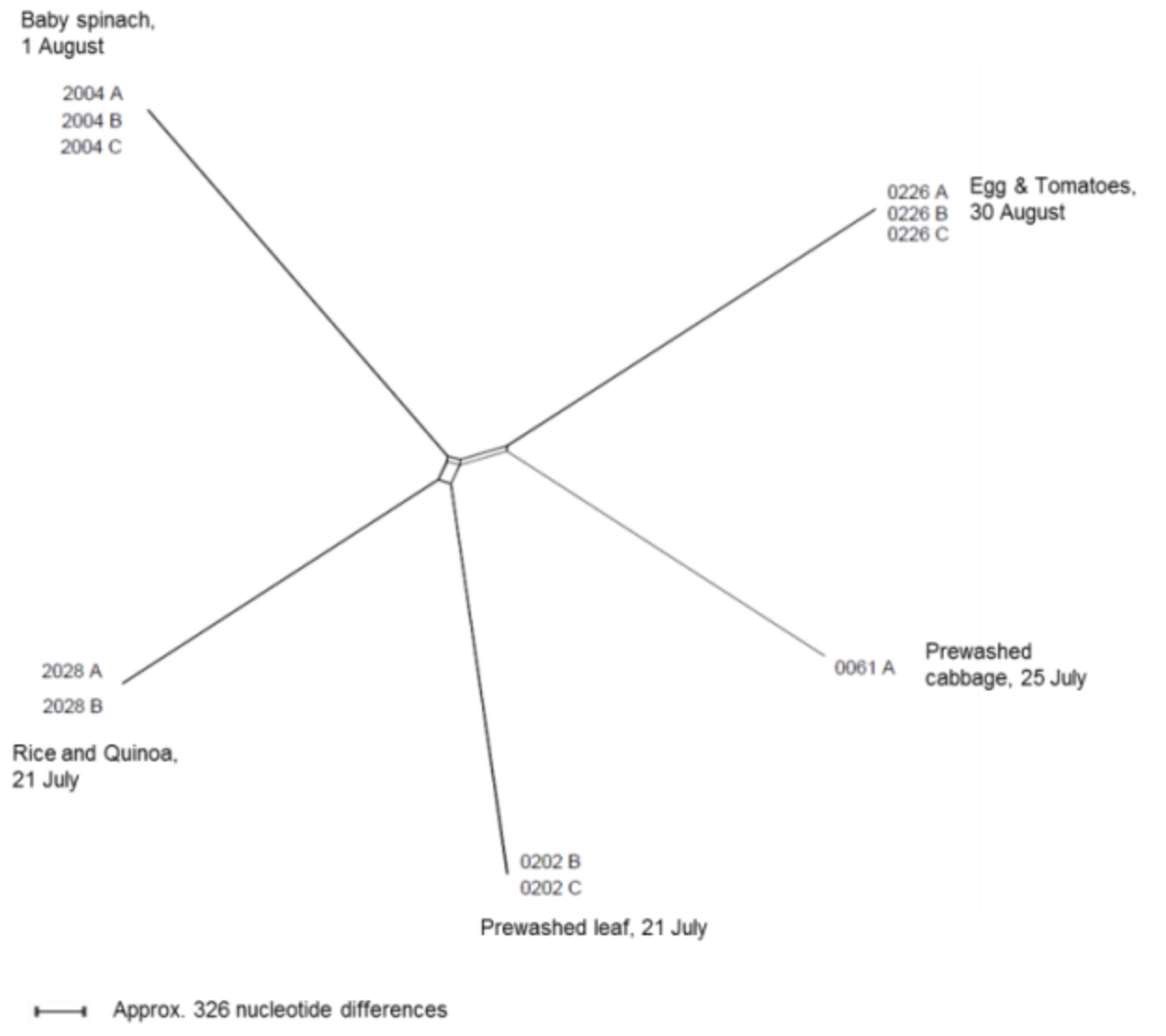
- For broader applications of high throughput sequencing, see “Next Generation Sequencing in the Food Industry”, Food Analysis Supplement 2016, New Food.



OriGen® Case Study

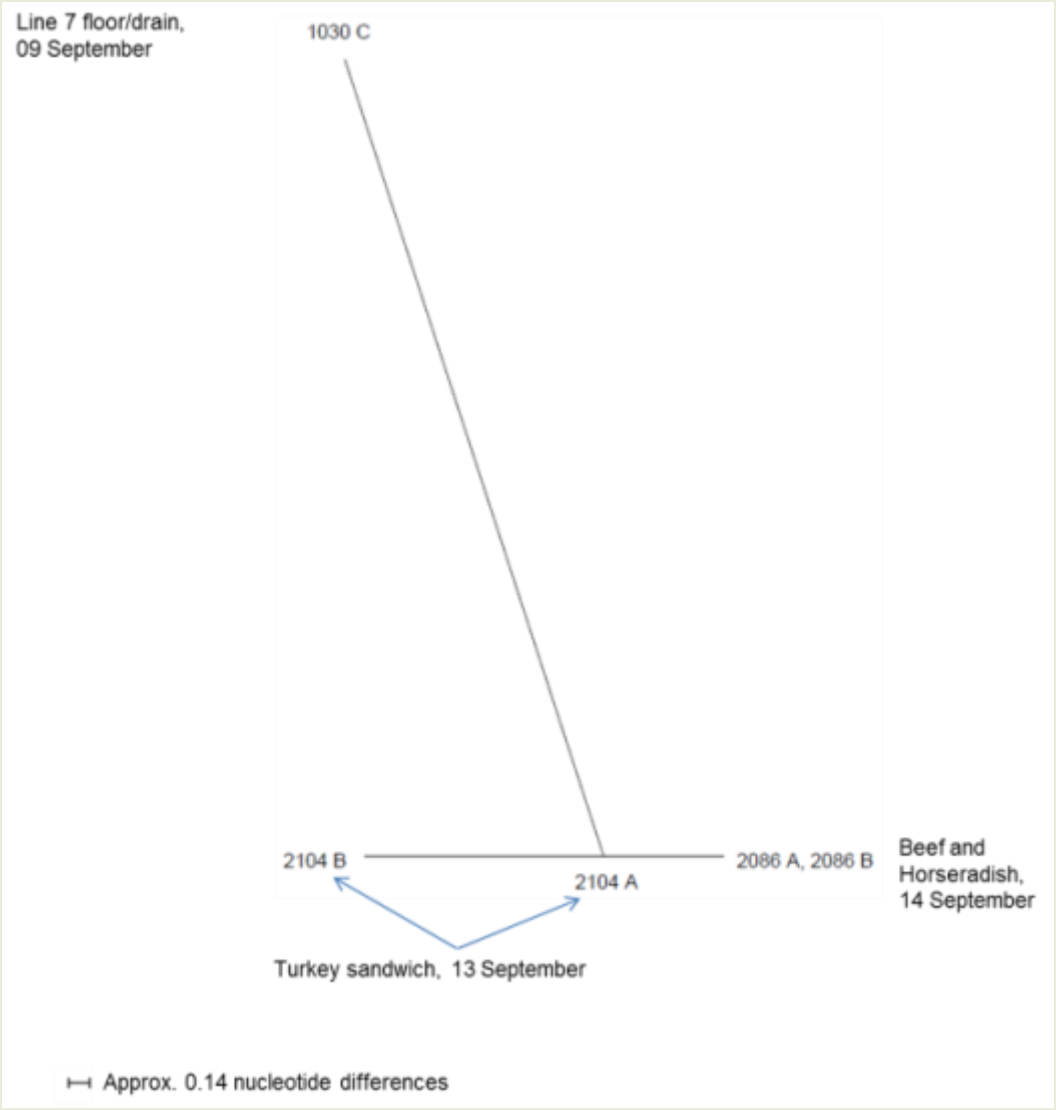
- Food manufacturer was looking to better understand *Listeria spp.* persistence vs sporadic recontamination in its facility, as well as any links to raw materials
- Contract lab performed routine swabbing, culturing and species ID of *Listeria*
- Fera obtained 42 cultures which after DNA extraction produced sequence data passing QC thresholds

Listeria seeligeri



- No evidence of common sources of *L. seeligeri* contamination was found among any of the sample types

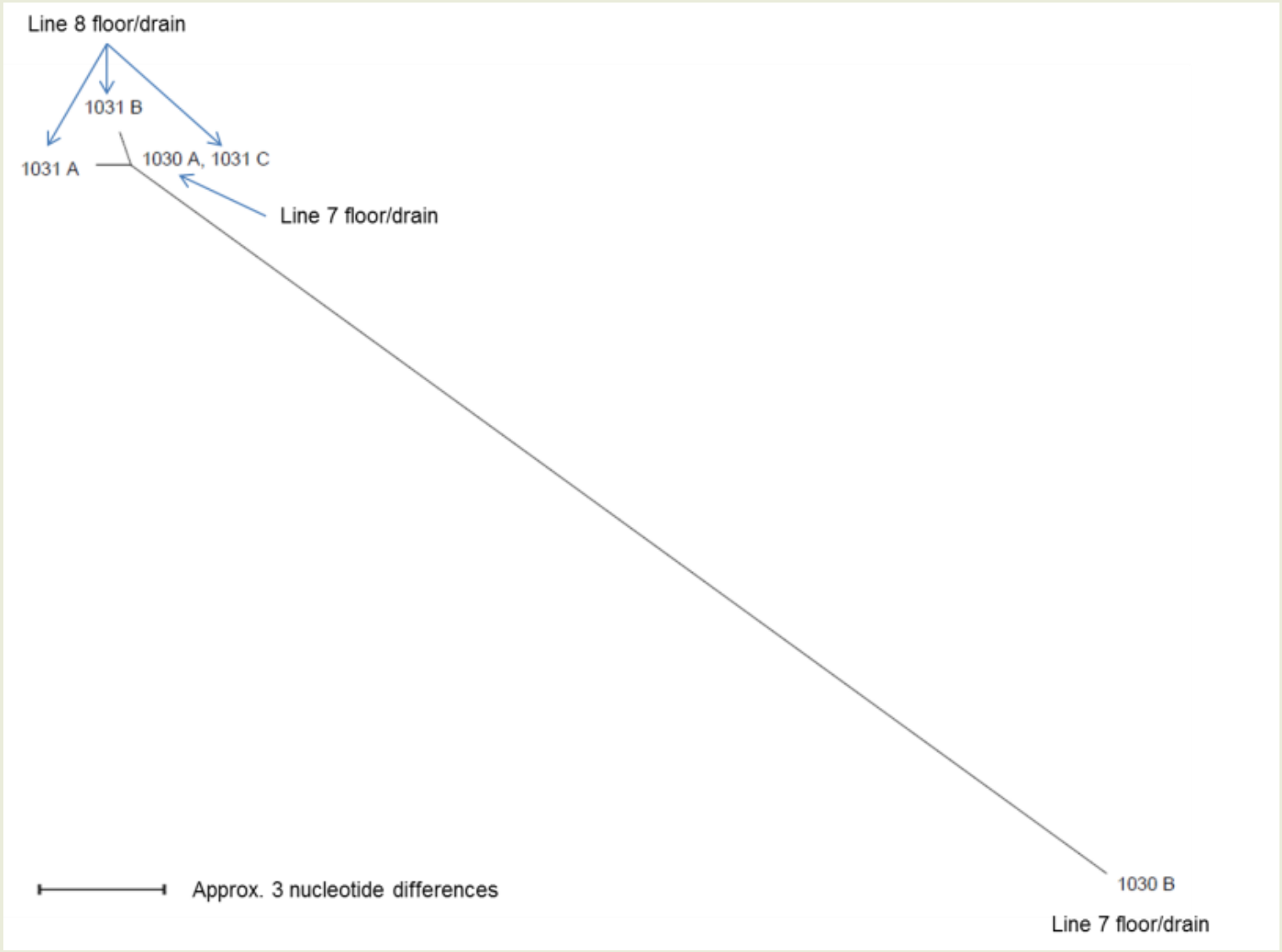
Listeria welshimeri



- All isolates of *L. welshimeri* were highly related, including those from a drain and a finished product

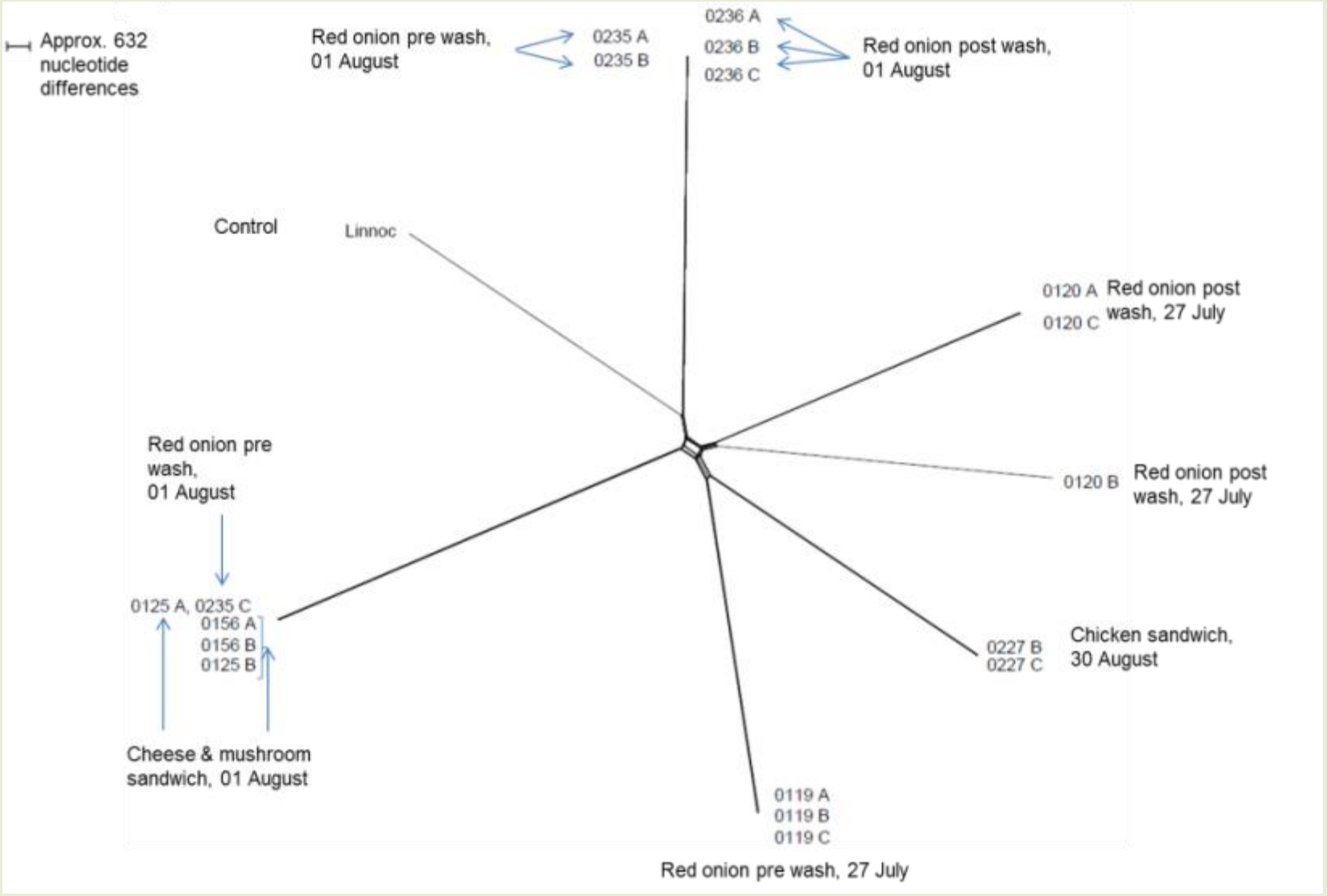
Listeria monocytogenes

Listeria monocytogenes magnified



- Two strains of *L. monocytogenes* were present, and samples from different production lines were highly related

Listeria innocua



- High diversity of *L. innocua* was observed in red onion samples
- Evidence of persistence of *L. innocua* from pre- to post-washing of red onions, and same type found in red onions and finished product

Food Metagenomics

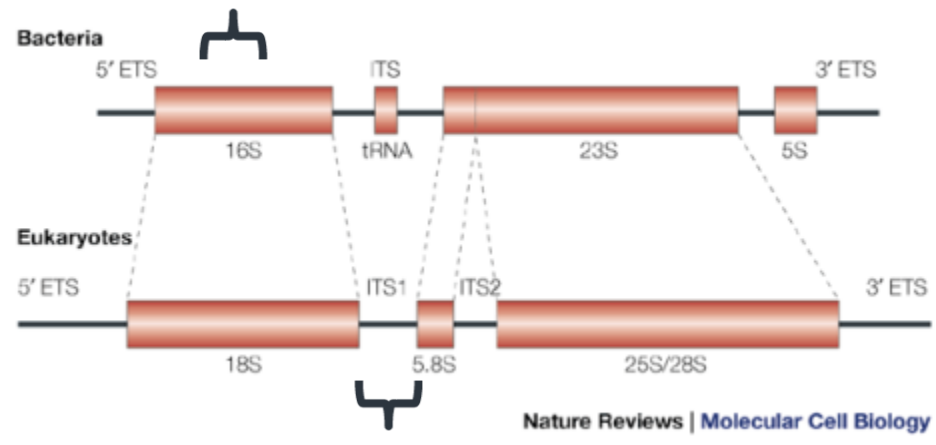
- Overall Project Aim: to develop a new non-targeted approach for determining origin and attribution of food products using microbial fingerprints

- Focussed on two foodstuffs, Pacific oysters (*Crassostrea gigas*) from the UK and France, and Stilton cheese

Metabarcoding

- DNA extracted from oyster gills/cheese samples
- Variable region V4 of 16S gene amplified using PCR for assessing bacterial populations in oysters and cheese
- Internal Transcribed Spacer (ITS) region 1 used for assessing fungal populations in cheese
- Products sequenced on illumina MiSeq platform

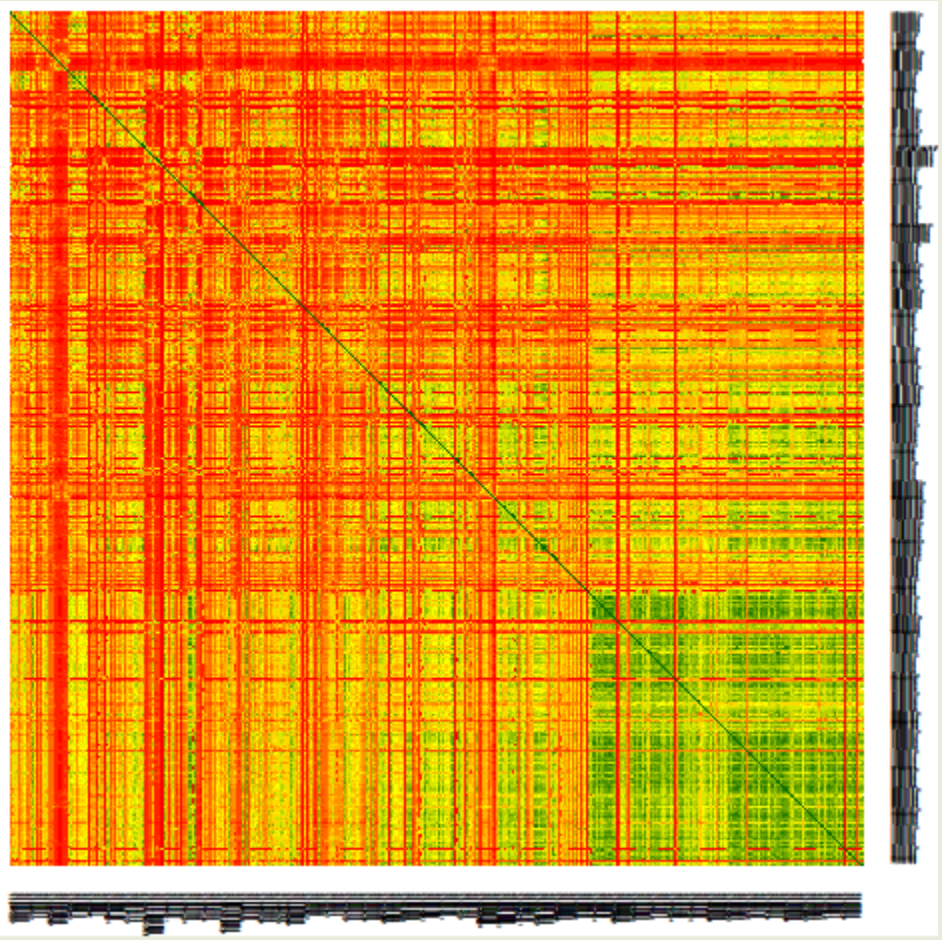
16S V4



ITS1

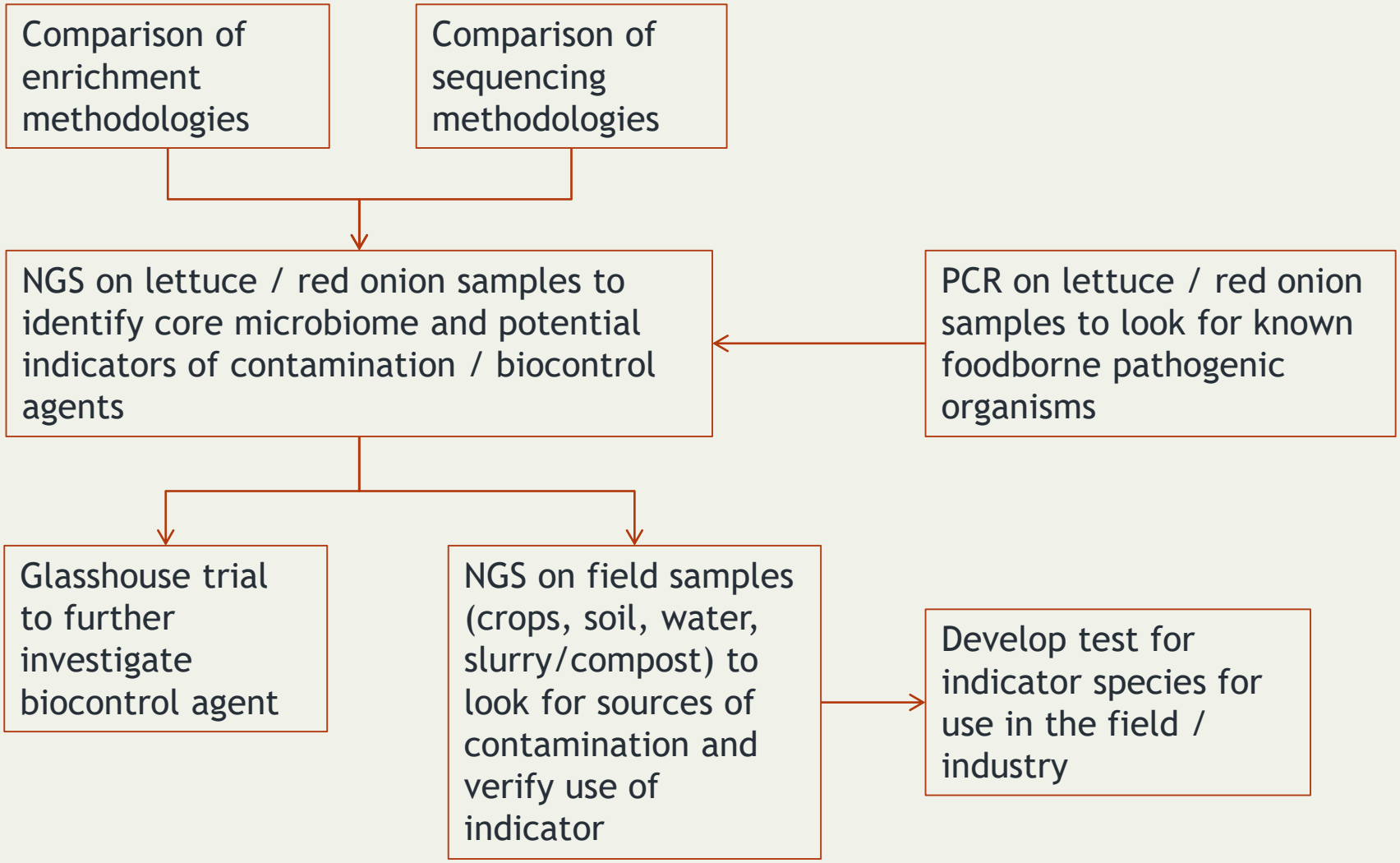
Lafontaine & Tollervey, 2001, doi:10.1038/35080045

Oyster results



- Heatmap showing Bray-Curtis dissimilarity metric. Samples sorted by month and year (L->R on x-axis).
- Some areas of green (similarity between sites, within time points), but many areas of red
- ANOSIM supports the conclusion that samples are too diverse to reliably assign to location
- Higher hopes for cheese

Non-targeted detection



Initial Results

Comparison of enrichment methodologies

When compared to poly-A capture, ribosomal depletion gave significantly more reads of:

- Higher quality
- Reads mapping to the internal extraction control
- Significantly higher Shannon-diversity indicating a greater detection of species (figure 1)
- Proportionally less reads belonging to Eukaryotes (figure 2)

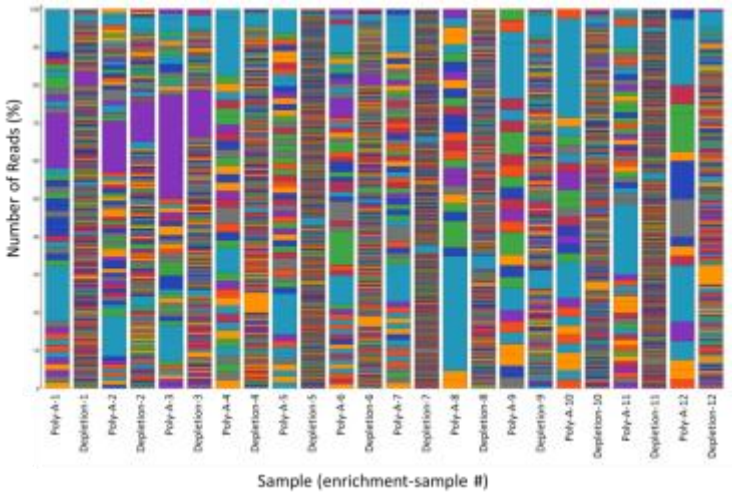


Figure 1: Graph showing the totalcomposition of each sample (1-12) for both enrichment methods - poly-A capture (Poly-A) and ribosomal depletion (Depletion) - as a percentage of reads obtained.

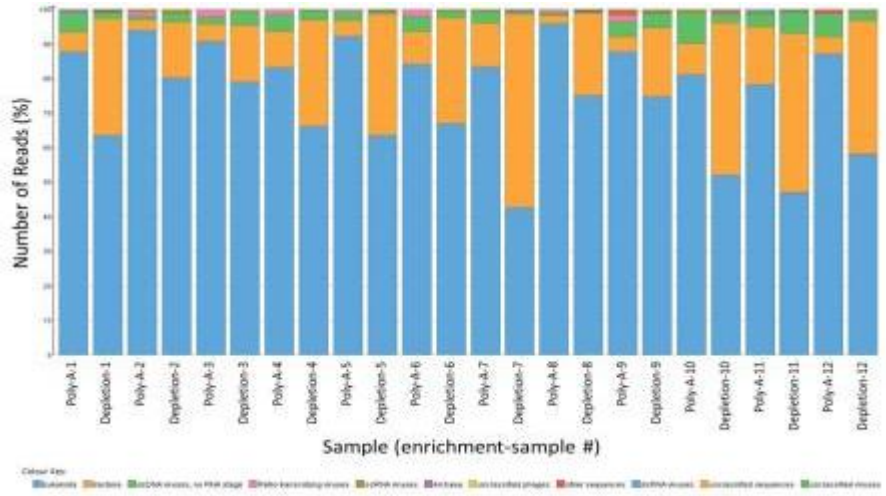


Figure 2: Graph showing the domain level composition of each sample (1-12) for both enrichment methods - poly-A capture (Poly-A) and ribosomal depletion (Depletion) - as a percentage of reads obtained.

Erin Lewis

Summary

- WGS being used to understand origins of microbial contamination of manufacturing facilities
- Metarbarcoding approach used to investigate authenticity and attribution of foodstuffs
- Non-targeted sequencing approach being trialled for detection of pathogens or indicator organisms

Acknowledgements

- Fera
  - Christine Conyers
  - Rachel Glover
  - Erin Lewis

- DEFRA/FSA/CEFAS - Development of Metagenomic Methods for Determination of Origin - phase 2
- FSA/PHE - NoVAS: Norovirus Attribution Study
- Newcastle University/IAFRI – Microbiological risk management tools for fruit and vegetable growers

- Contact: Edward Haynes
  - edward.haynes@fera.co.uk