Adoption of new and modified technologies: future proofing our experiments

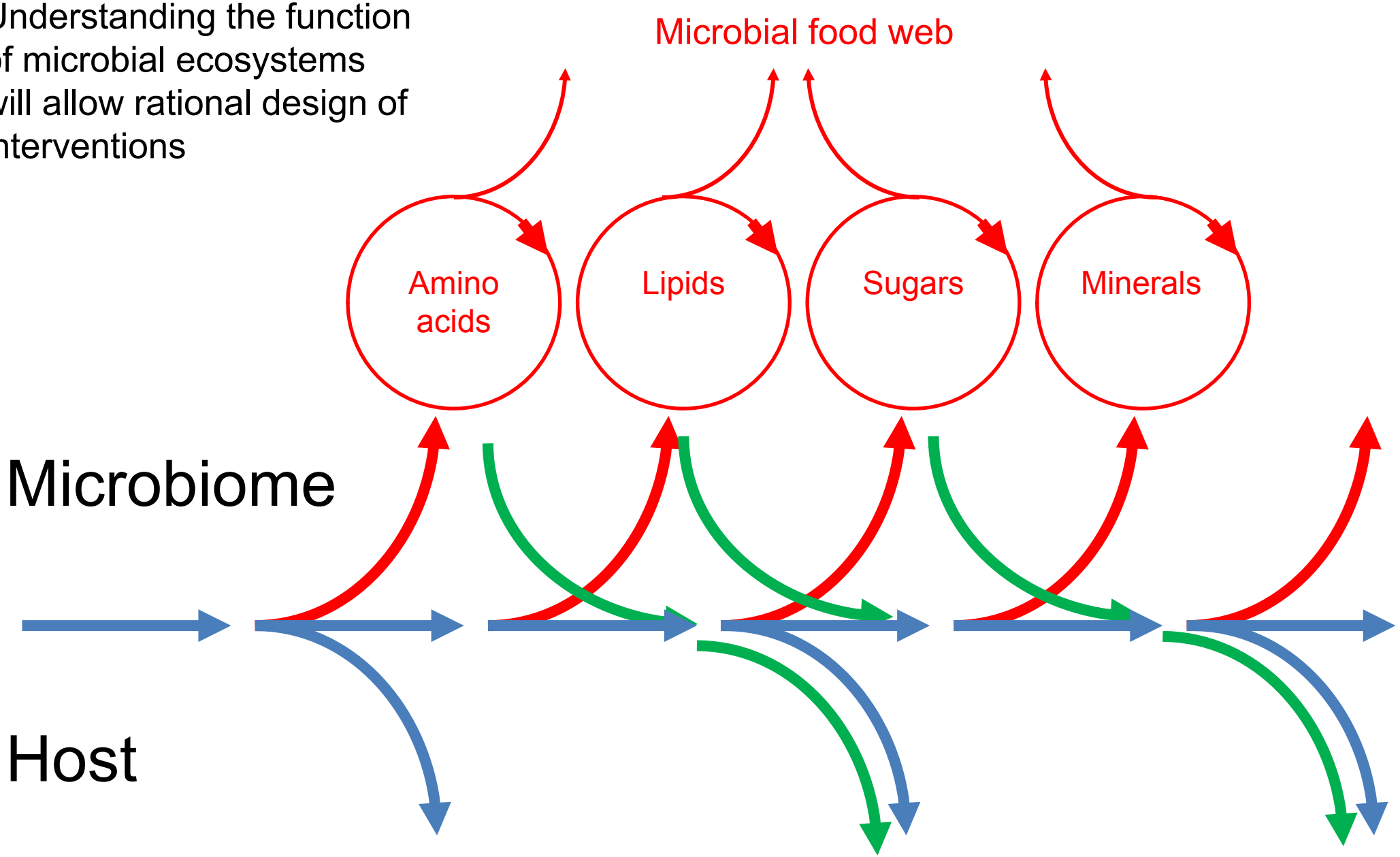
Mick Bailey
Professor of Comparative Immunology
Bristol University, UK

- This arose from a discussions at a meeting in London attended by Claire, Jordi and me.
- It was apparent that new technologies are continuously developing for:
  - DNA extraction
  - Amplification
  - Sequencing
  - Community standards
- However, it was also apparent that there is value in standardising on a single pipeline and accumulating a large, internally comparable database.
- The Cargill experience – around 20,000 global chicken faecal samples.

Standardise on old methods or use cutting edge technology?

At what stage should we adopt new technologies?

Understanding the function of microbial ecosystems will allow rational design of interventions

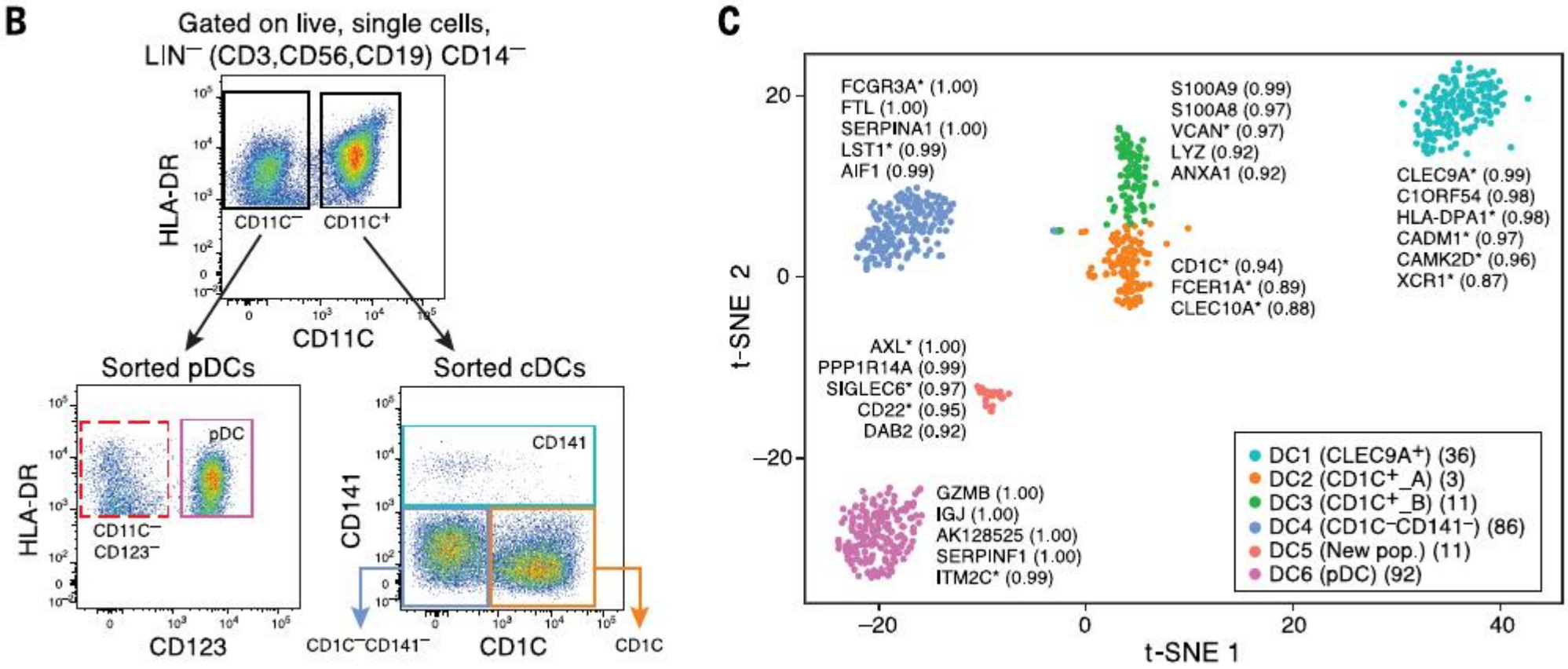


We are probably at least twenty years from this depth of understanding

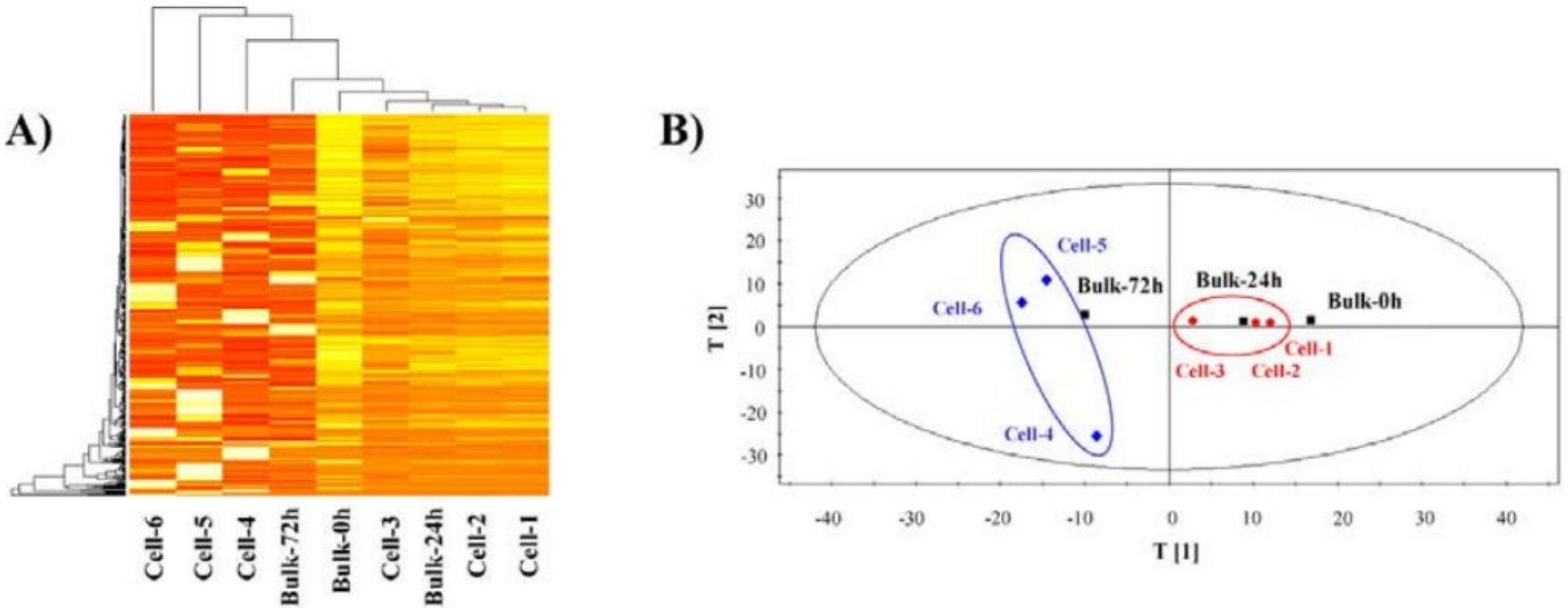
What do we actually want to know?

• Whether there are differences between individuals in composition of the microbiome?	DGGE
• The structure of the intestinal microbiome?	16S rRNA
• The metabolic pathways which the intestinal microbiome can use?	Metagenomics
• The functions of the intestinal microbial ecosystem?	Microbial transcriptomics, proteomics, metabolomics
• The ways in which the microbial ecosystem interacts with the host	MALDI-imaging, single-cell and Laser-Capture transcriptomics

Single cell RNAseq

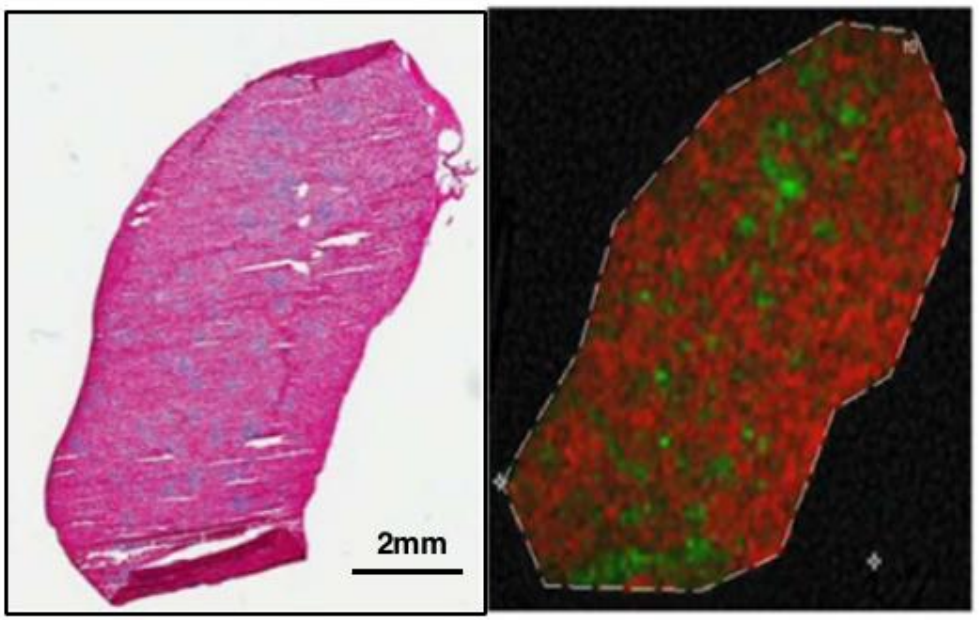


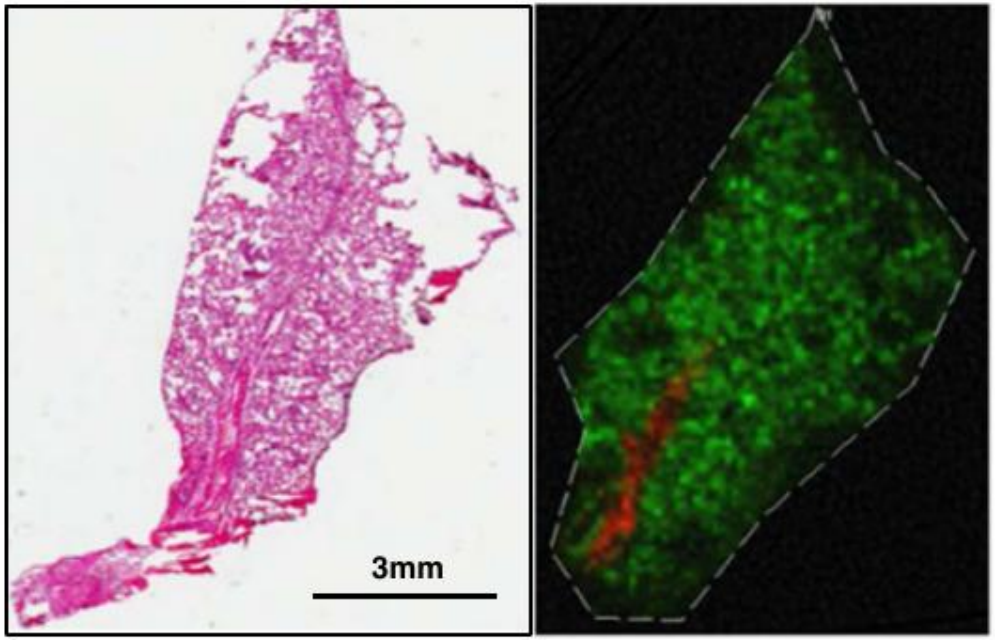
Villani et al., Science 356, 283 (2017) 21 April 2017



Chen et al. RNA-seq based transcriptomic analysis of single bacterial cells
Article in Integrative Biology · September 2015
DOI: 10.1039/c5ib00191a

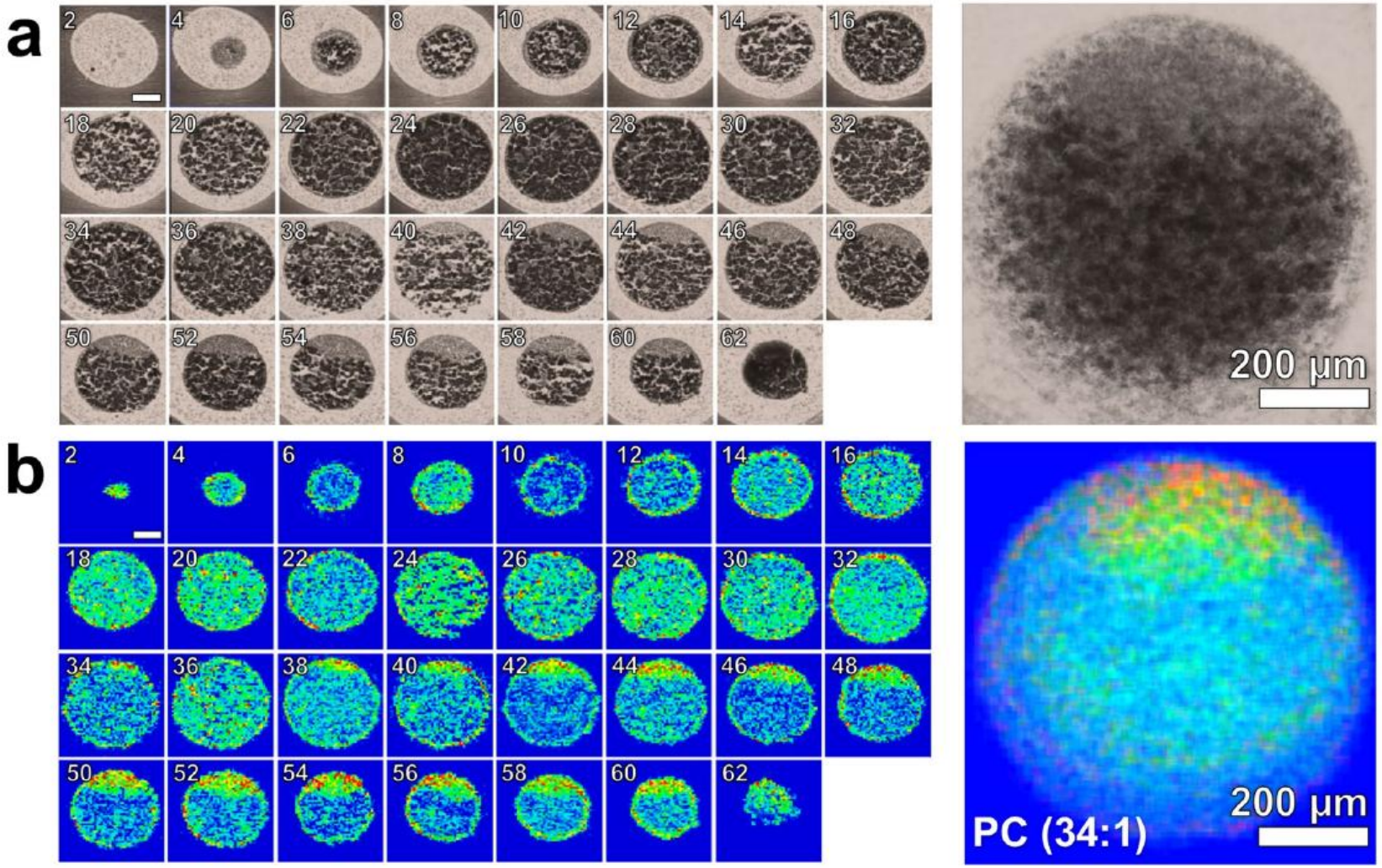
MALDI-imaging: metabolites and peptides at 10-20μ resolution



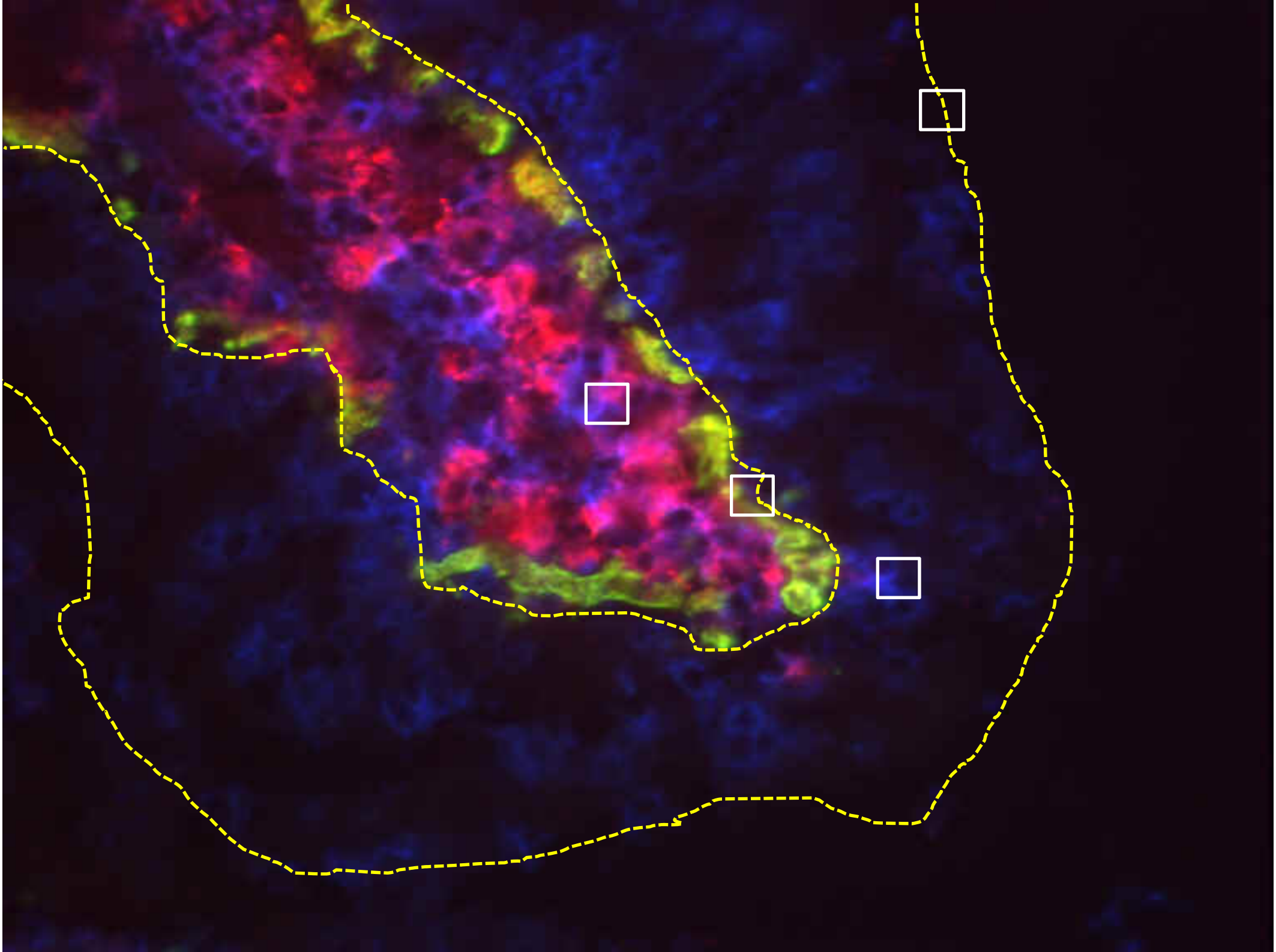


Cazares et al. BMC Microbiology (2015) 15:101
DOI 10.1186/s12866-015-0431-7

MALDI-imaging: metabolites and peptides at 10-20μ resolution



Duenas et al, 2017. Scientific Reports 7:14946. DOI:10.1038/s41598-017-14949-x



RESEARCH ARTICLE

Sewage Reflects the Microbiomes of Human Populations

Ryan J. Newton,[a] Sandra L. McLellan,[a] Deborah K. Dila,[a] Joseph H. Vineis,[b] Hilary G. Morrison,[b] A. Murat Eren,[b] Mitchell L. Sogin[b]

School of Freshwater Sciences, University of Wisconsin—Milwaukee, Milwaukee, Wisconsin, USA[a]; Josephine Bay Paul Center, Marine Biological Laboratory, Woods Hole, Massachusetts, USA[b]

ABSTRACT Molecular characterizations of the gut microbiome from individual human stool samples have identified community patterns that correlate with age, disease, diet, and other human characteristics, but resources for marker gene studies that consider microbiome trends among human populations scale with the number of individuals sampled from each population. As an alternative strategy for sampling populations, we examined whether sewage accurately reflects the microbial community of a mixture of stool samples. We used oligotyping of high-throughput 16S rRNA gene sequence data to compare the bacterial distribution in a stool data set to a sewage influent data set from 71 U.S. cities. On average, only 15% of sewage sample sequence reads were attributed to human fecal origin, but sewage recaptured most (97%) human fecal oligotypes. The most common oligotypes in stool matched the most common and abundant in sewage. After informatically separating sequences of human fecal origin, sewage samples exhibited ~3× greater diversity than stool samples. Comparisons among municipal sewage communities revealed the ubiquitous and abundant occurrence of 27 human fecal oligotypes, representing an apparent core set of organisms in U.S. populations. The fecal community variability among U.S. populations was significantly lower than among individuals. It clustered into three primary community structures distinguished by oligotypes from either: *Bacteroidaceae*, *Prevotellaceae*, or *Lachnospiraceae/Ruminococcaceae*. These distribution patterns reflected human population variation and predicted whether samples represented lean or obese populations with 81 to 89% accuracy. Our findings demonstrate that sewage represents the fecal microbial community of human populations and captures population-level traits of the human microbiome.

Can we use slurry handling systems to monitor farm-level microbiota (within- and between-farm variation) and relate this to performance and health?

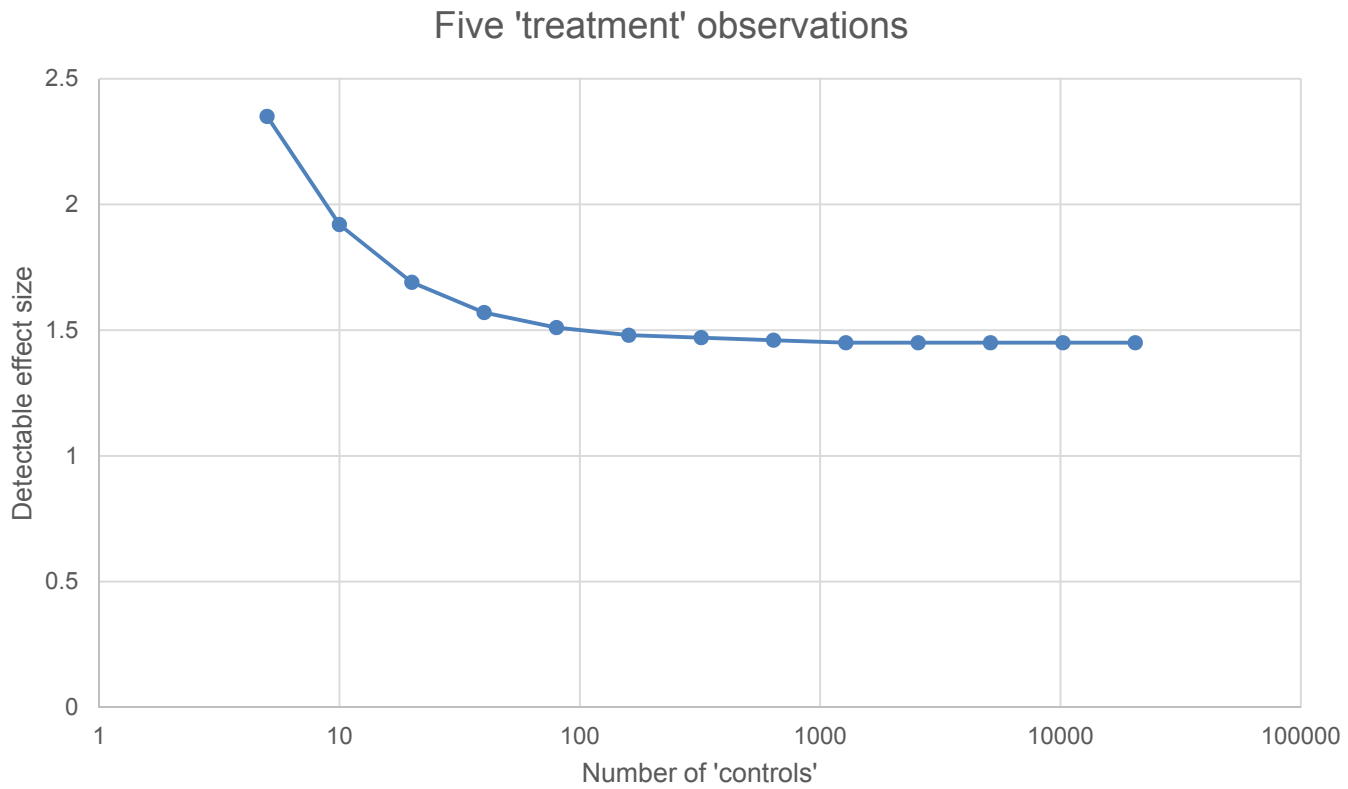
Standardise on old methods or use cutting edge technology?

- All of the ‘omics technologies, but particularly sequencing, are developing much faster than any techniques for data analysis
- If we don’t move to new technologies, we limit our ability to gain valuable information
- The simple answer is that we should always use the most cutting edge technology
- However, there are obvious reasons why that may not be the best strategy

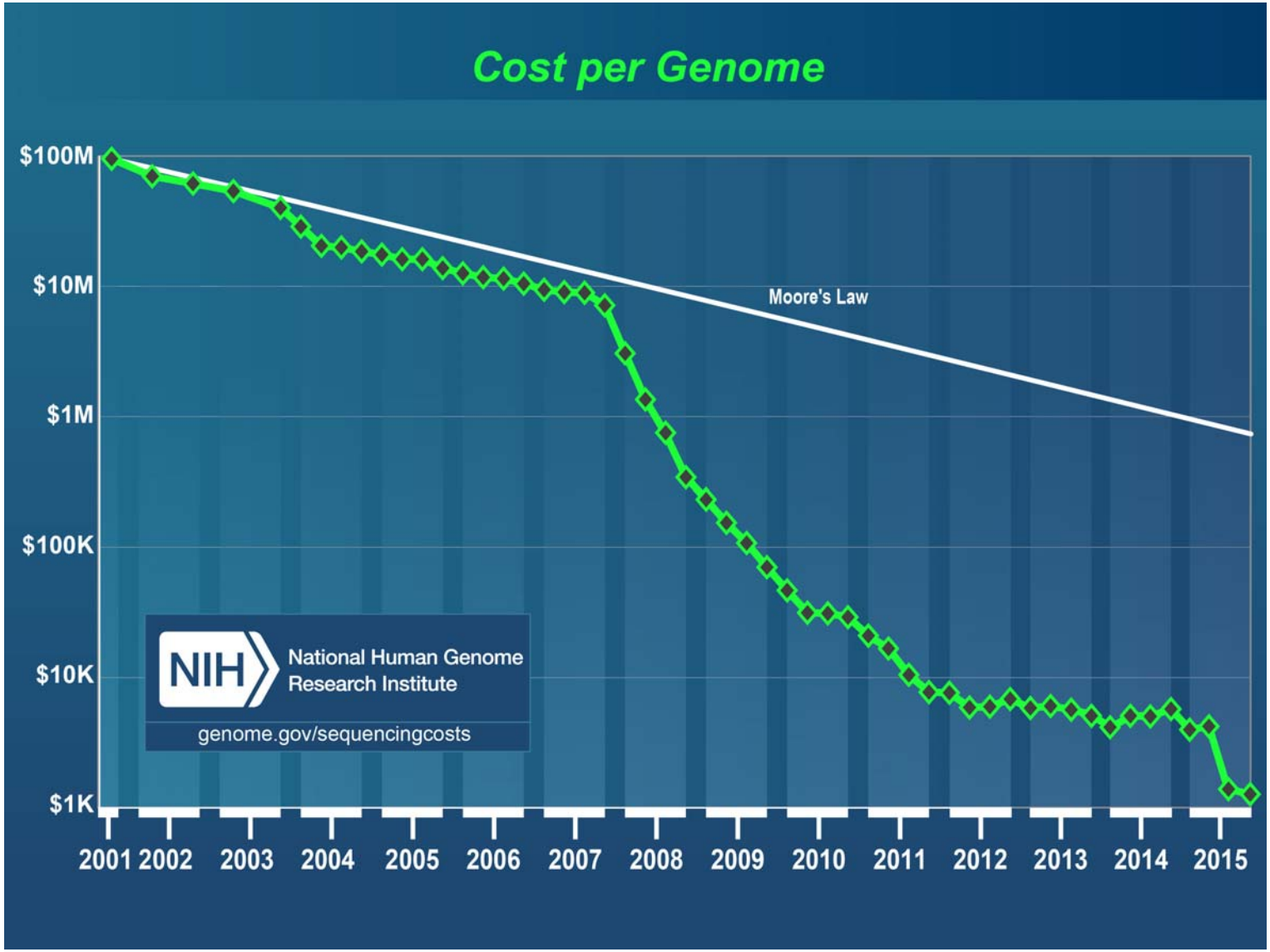
Standardise on old methods or use cutting edge technology?

- Old technology is often cheaper until the costs of the new technologies drop
- Each time we move to a new technology, we make much of our previous results obsolete.
- We lose the ability to compare our new samples with our previous samples
- This is a simple statistical power calculation – the larger the ‘control’ dataset, the fewer new observations are needed to identify an effect

- Increasing the number of ‘controls’ means an experiment can detect smaller effect sizes.
- Or that an effect can be detected with fewer ‘treated’ animals.
- …although a fully balanced design is still more powerful
- …and the value of increasing ‘controls’ shows diminishing returns



- Whether laboratories engage with standardisation depends on the value they perceive that it has :
  - How much information do we gain by using the latest, cutting edge technology?
  - How much information do we gain by using large data sets obtained by standardising technology over a long period of time?
- Can we use information theory or statistical power calculations to estimate both of these? The reduction in entropy, or in the error term, associated with each approach?

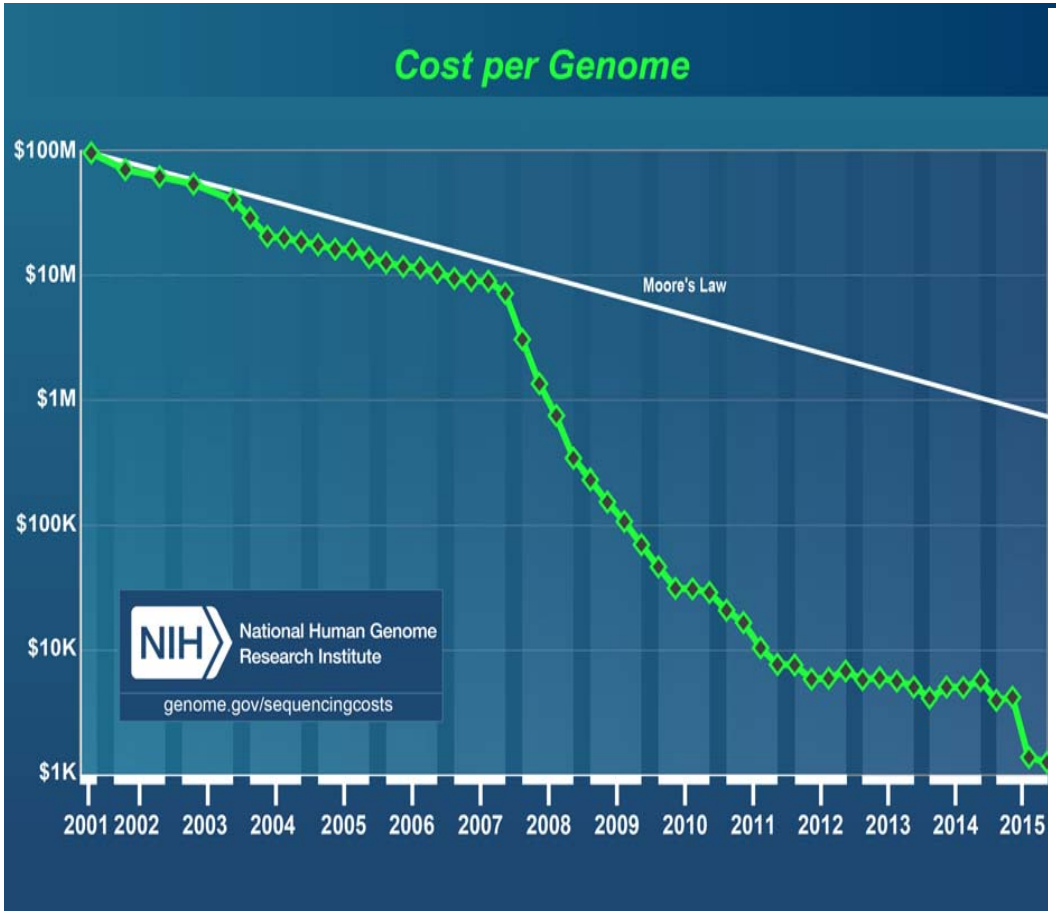


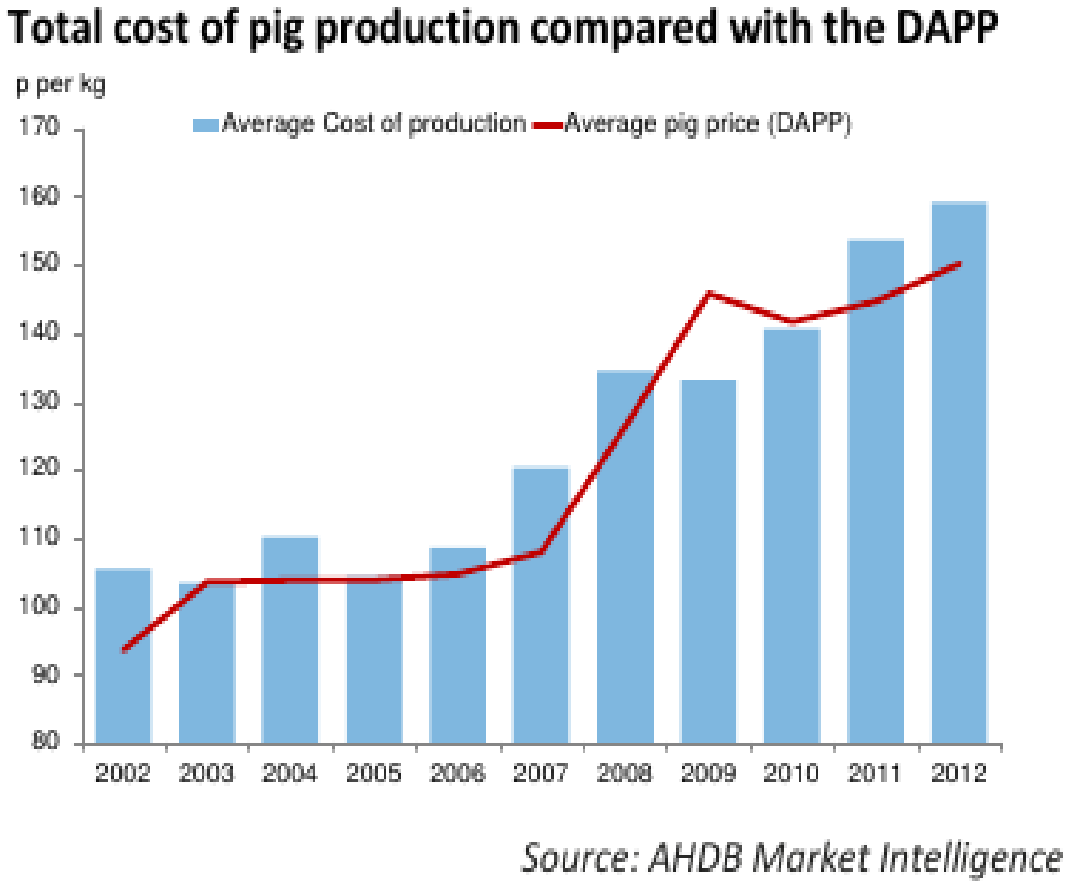
- Costs of sequencing are decreasing faster than our ability to analyse using existing algorithms

- Unless progress in development of ‘omics technologies slows, we will continue to outpace our ability to properly analyse the data which we can generate
- We need to develop new, smarter algorithms to make better use of existing computing power, which may use fuzzier logic (with inevitable loss of precision)
- We need to engage with advances in computing power through our institutional research groups (what are the likely real benefits of quantum computing?)

- How much should we worry about making our sequence data comparable?
- Sequencing will get cheaper and cheaper
- Animal experiments will not

Costs of sequencing and analysis are decreasing, costs of animals are increasing





It makes a lot more sense to archive valuable samples with good metadata in accessible biobanks

- Ethical restrictions on animal experiments are also making them more and more expensive
- In the UK, we already have to convince our ethical review boards that similar experiments have not already been carried out elsewhere
- We may not be able to convince our ethical review processes that we need to keep repeating experiments to take new, different samples
- Maximising the future value of animal experiments supports the principles of ‘Reduction, Refinement and Replacement’

- New techniques which we will have access to in the near future are already visible
- We should be planning our sample collection, replication, storage and archiving with new technologies in mind
- If we do, we minimise the problem of when to switch – we can re-analyse existing, well-archived samples with new technologies
- But this needs planning into applications for funding:
  - Horizon-scanning for new technologies
  - Technical support for taking samples
  - Routinely splitting samples (within and between repositories)
  - Formal, standardised archiving of metadata
  - Long-term, fault-tolerant storage