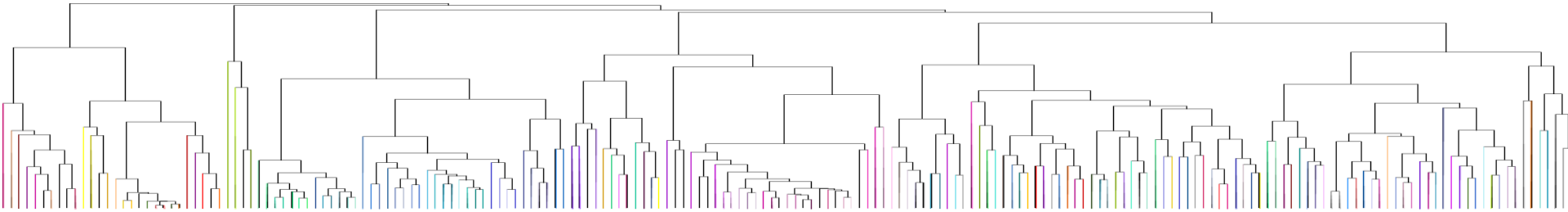
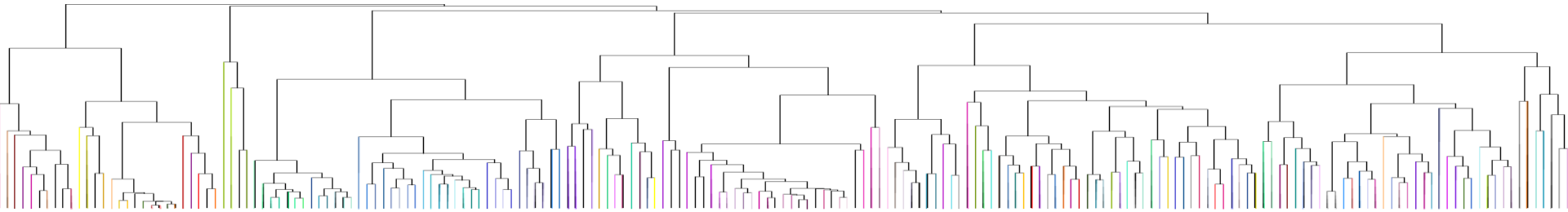


Library preparation for 16S rRNA sequencing



Shantelle Claassen-Weitz

Division of Medical Microbiology

Department of Pathology

tellafiela@gmail.com

University of Cape Town • iYunivesithi yaseKapa • Universiteit van Kaapstad

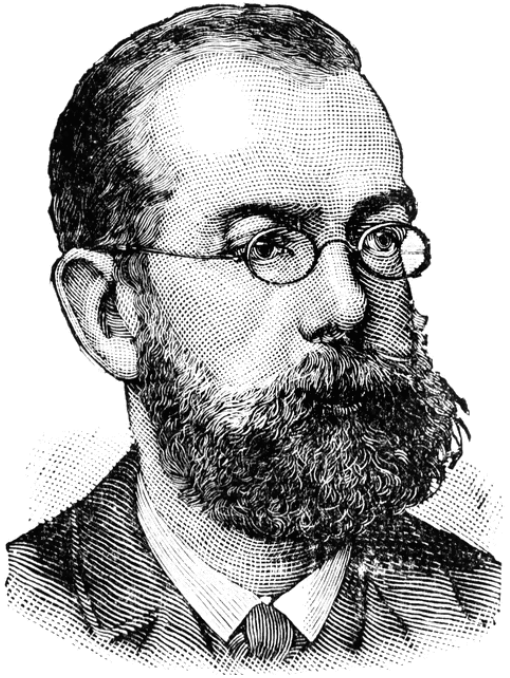
SPES BONA

DRAKENSTEIN CHILD LUNG HEALTH STUDY

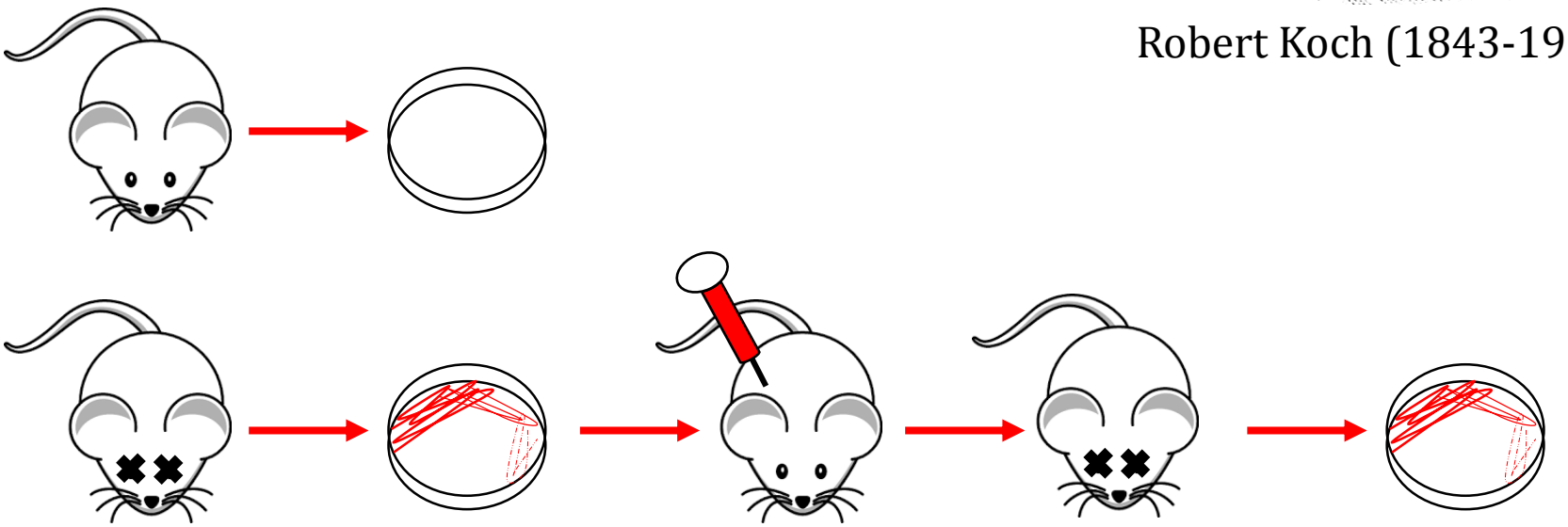
The “golden age of microbiology”

The schools of Robert Koch and Louis Pasteur identified **pathogenic** microorganisms as the causative agents for many communicable human diseases.

The Koch’s postulates were essential for demonstrating that a **specific microbe** isolated in pure culture **from a lesion of the disease could produce illness in a healthy animal**, which gave rise to the concept of **pathogenic microbes**.



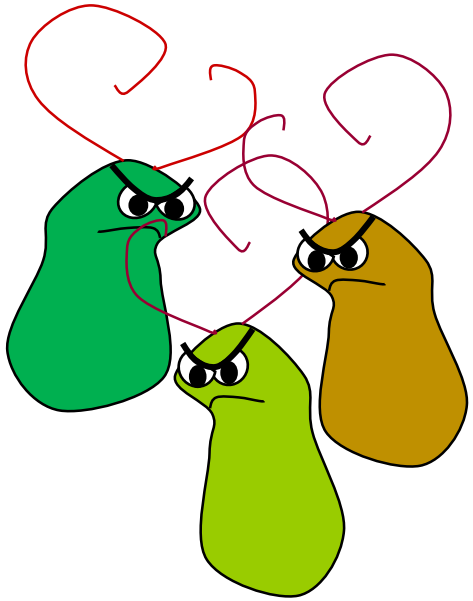
Robert Koch (1843-1910)



The microbial community and health and disease

Recent advances in high-throughput sequencing technology has provided information on the **human microbiome** and its **physiological potential**.
For example, novel links between the gut microbiota, the largest microbial community inhabiting our body, and a wide variety of non-communicable diseases, including arthritis, arteriosclerosis, obesity, diabetes, colorectal cancer, liver cancer, asthma and neurodevelopmental disorders, have been reported.

Dysbiosis:
where an **imbalance in the microbial community** is associated with **disease conditions**



Bloom of pathobionts

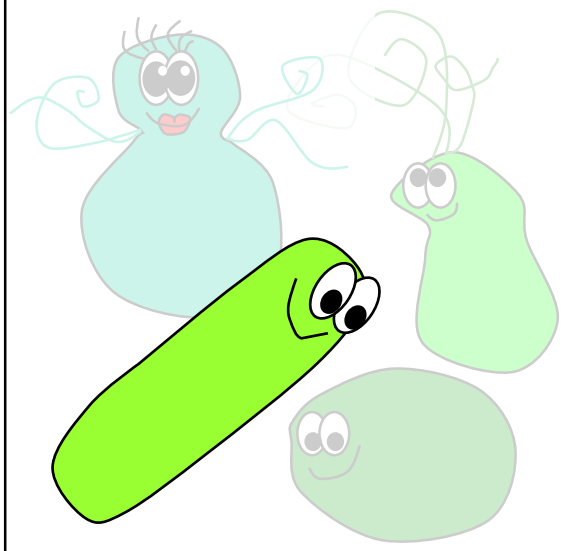
- An overgrowth of commensal microbiota that have the potential to cause pathology.
- An example of such population expansion is the outgrowth of the bacterial family Enterobacteriaceae, which is frequently observed in enteric infection and inflammation.

Chow & Mazmanian (2010)., *Cell Host Microbe,* **7**, pages 265–276.
Stecher, Maier. & Hardt. (2013)., *Nat. Rev. Microbio*l. **11**, pages 277–284.
Frank et al. (2007)., *Proc. Natl Acad. Sci. USA,* **104**, pages 13780–13785.; Garrett et al. (2007)., *Cell,* **131**, pages 33–45.

The microbial community and health and disease

Recent advances in high-throughput sequencing technology has provided information on the **human microbiome** and its **physiological potential**.
For example, novel links between the gut microbiota, the largest microbial community inhabiting our body, and a wide variety of non-communicable diseases, including arthritis, arteriosclerosis, obesity, diabetes, colorectal cancer, liver cancer, asthma and neurodevelopmental disorders, have been reported.

Dysbiosis:
where an **imbalance in the microbial community** is associated with **disease conditions**



Loss of commensals

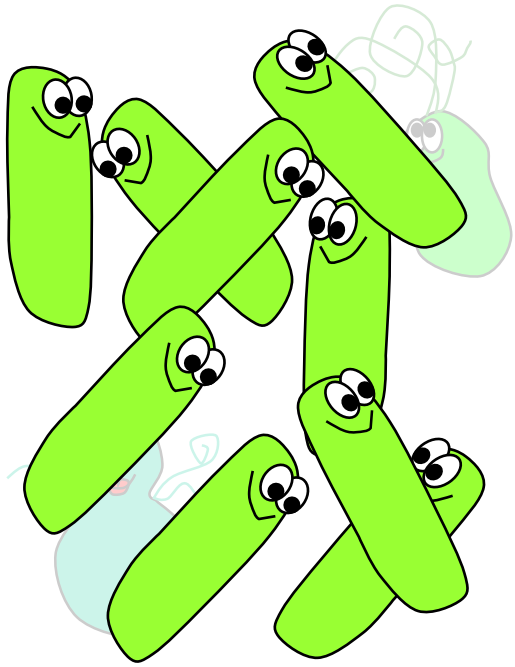
- Dysbiosis frequently features the reduction or complete loss of normally residing members of the microbiota.
- Replenishment of diminished commensal bacteria has also proved effective against enteric infection

Korem et al. (2015)., Science, 349, pages 1101–1106.
Buffington et al. (2016)., Cell, 165, pages 1762–1775.
Hsiao et al. (2013)., Cell, 155, pages 1451–1463.
Buffie, et al. (2015)., Nature, 517, pages 205–208.

The microbial community and health and disease

Recent advances in high-throughput sequencing technology has provided information on the **human microbiome** and its **physiological potential**.
For example, novel links between the gut microbiota, the largest microbial community inhabiting our body, and a wide variety of non-communicable diseases, including arthritis, arteriosclerosis, obesity, diabetes, colorectal cancer, liver cancer, asthma and neurodevelopmental disorders, have been reported.

Dysbiosis:
where an **imbalance in the microbial community** is associated with **disease conditions**

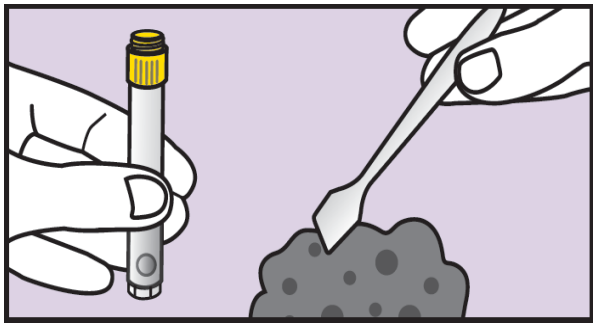


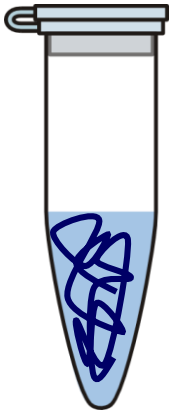
Loss of diversity

- Low bacterial diversity has been documented in the context of dysbiosis induced by abnormal dietary composition, IBD, AIDS and type 1 diabetes (T1D), among many other conditions.

Cotillard et al. (2013). *Nature,* **500**, pages 585–588.
Le Chatelier et al. (2013). *Nature,* **500**, pages 541–546.
Mosca, Leclerc, & Hugot. (2016)., *Front. Microbiol*. **7**, 455.

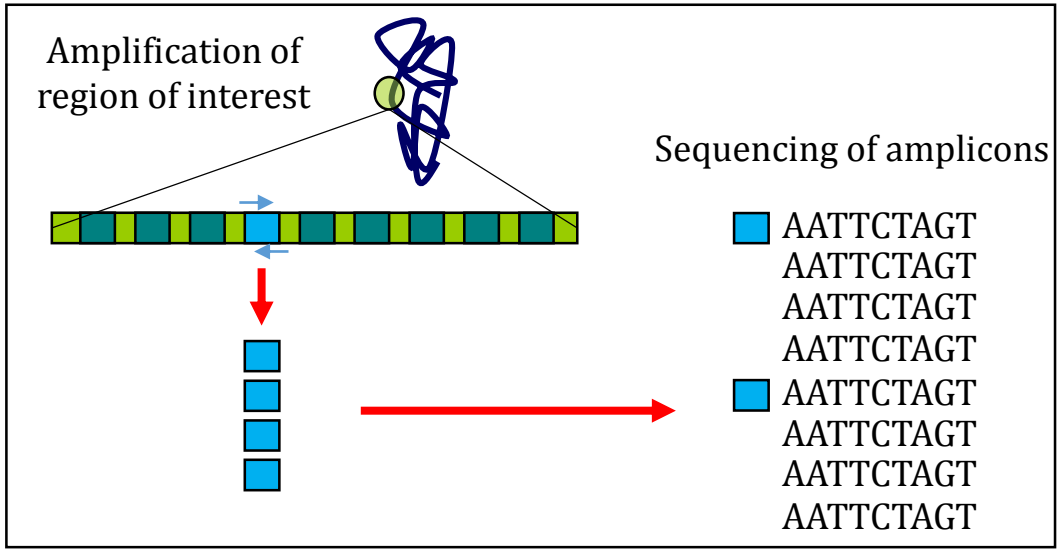
How do we study our microbial communities?



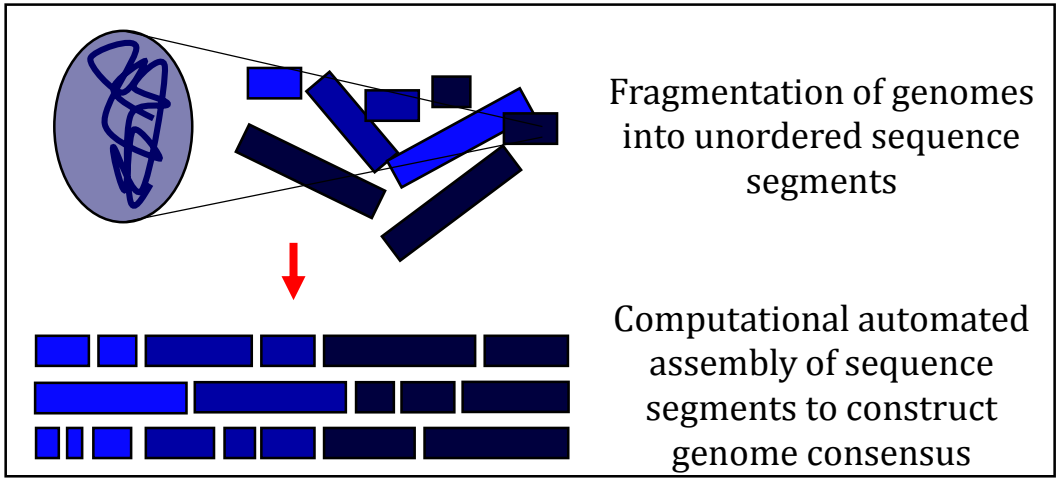


Nucleic acid

16S rRNA sequencing

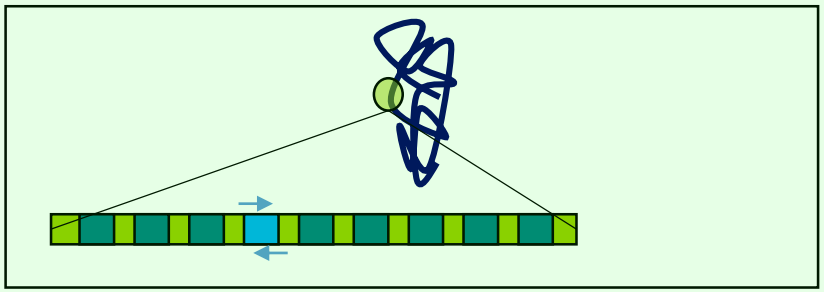


Shotgun metagenomics sequencing



Benefits and limitations of the two methods used for microbial community analysis

16S rRNA sequencing



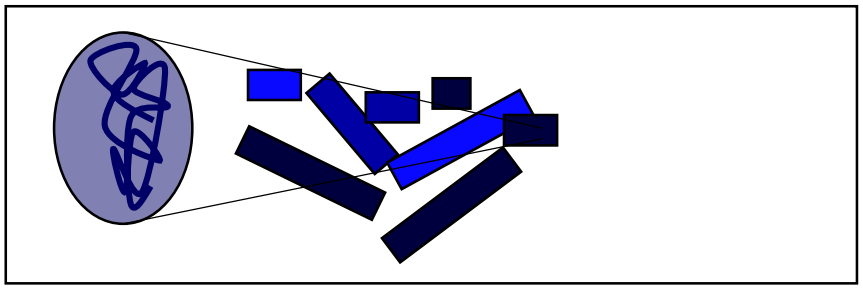
Benefits:

1. cost effective
2. data analysis can be performed by established pipelines
3. a large body of archived data is available for reference.

Limitations:

1. sequences only a single region of the bacterial genome
2. classifications often lack accuracy at the species level
3. specific genes are not directly sequenced, but rather predicted based on the OTUs

Shotgun metagenomics sequencing



Benefits:

1. sequence broad regions of the genome
2. identifies approximately twice as many species
3. identify organisms in additional kingdoms including viruses, fungi and protozoa
4. increased prediction of genes and functional pathways

Limitations

1. expensive
2. computationally intensive

Ranjan et al. (2016).
Biochem Biophys Res Commun. 469(4): 967–977.

Addressing the “fine print” behind 16S rRNA sequencing when constructing an amplicon library

Addressing the "fine print" behind 16S rRNA sequencing when constructing an amplicon library

- **Sample collection:**
  Sample collection could influence the microbiota profile obtained following sequencing

- **DNA isolation:**
  Template concentration - may significantly impact on sample profile variability for most samples.
  Template extraction method – may influence the bacterial composition extracted from the sample (as well as the purity and DNA yield).

- **PCR amplification**
  PCR bias and inhibitors – PCR bias will affect downstream sequencing results
  Amplification of contaminants – from extraction kits and PCR reagents amongst others

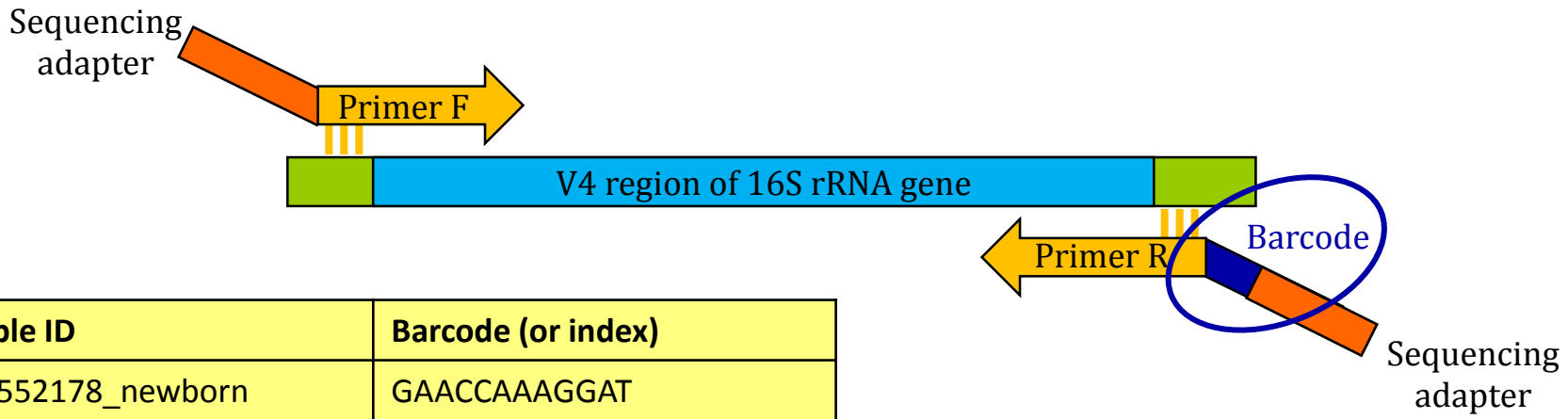
Kennedy et al., (2014), *App Environ Microbiol*, **80**, 5717-5722; Kennedy et al., 2014, *PloS One*, **9**, e88982, Weiss et al., (2014), *Gen Biol*, **15**, 564; Patel et al., (2012), *PloS One*, **7**, e30619; Jervis-Bardy, (2015), *Microbiome*, **3**, 19

Steps involved:

Step 1
Design your experiment

How many samples can be included in a sequencing run?

By using barcoded primers, numerous samples can be sequenced simultaneously (this is referred to as multiplexing).



Sample ID	Barcode (or index)
GITB552178_newborn	GAACCAAAGGAT
GITB522968_one_month	TACACGATCTAC
GITB536987_two_months	GCGATATATCGC
GITM528796_birth	CAGTGCATATGC
GITM564789_birth	TCCAAAGTGTTC
GITM566987_birth	GGCCACGTAGTA

Steps involved:

Step 1
Design your experiment

How many samples can be included in a sequencing run?

- The more samples included, the more cost effective the run, however with reduced sequencing depth.
- We typically process up to 384 reaction per run (4x96 well plates)
- It is critical to have a **"library prep manifest"** to document the positions of each sample together with its barcode and any other clinical or experimental data.

Sample ID	Mother or infant	DNA yield (ng/ul)	Plate position	Barcode (or index)
GITB552178_newborn	infant	25	P1A01	GAACCAAAGGAT
GITB522968_one_month	infant	39	P1A02	TACACGATCTAC
GITB536987_two_months	infant	60	P1A03	GCGATATATCGC
GITM528796_birth	mother	221	P1A04	CAGTGCATATGC
GITM564789_birth	mother	157	P1A05	TCCAAAGTGTTC
GITM566987_birth	mother	195	P1A06	GGCCACGTAGTA

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Plate 1

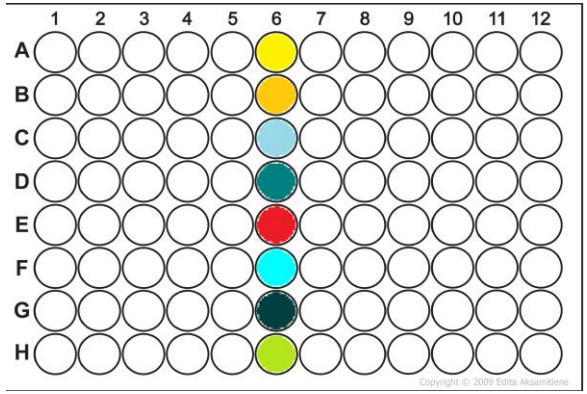
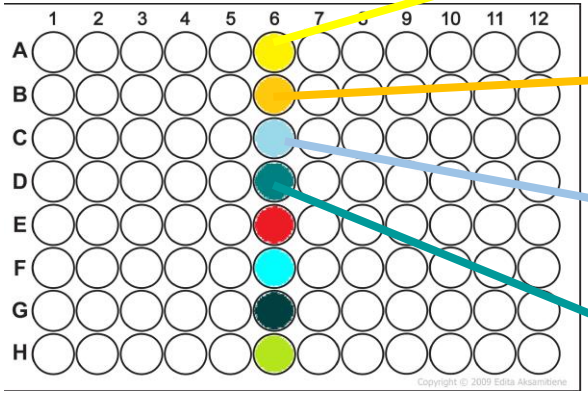


Plate 2



Between-run repeat (process any specimen in duplicate per run to measure reproducibility across runs)

Within-run repeat (process any specimen in duplicate on each of the plates to measure reproducibility)

Water used during PCR (to determine if any contamination was introduced during PCR reactions)

Water spiked with known bacterial DNA (to control for contaminants introduced during PCR reactions)

Plate 3

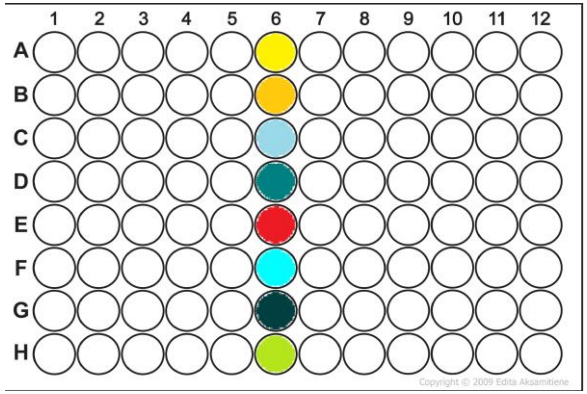
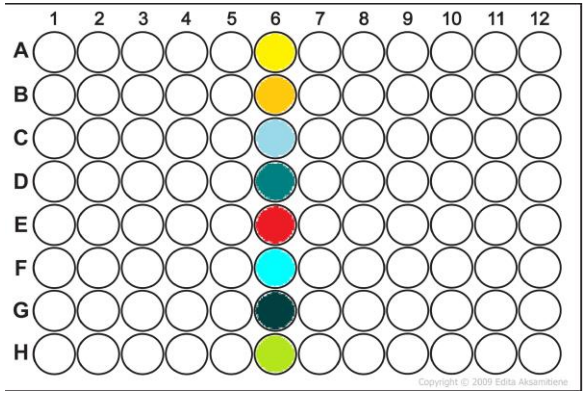


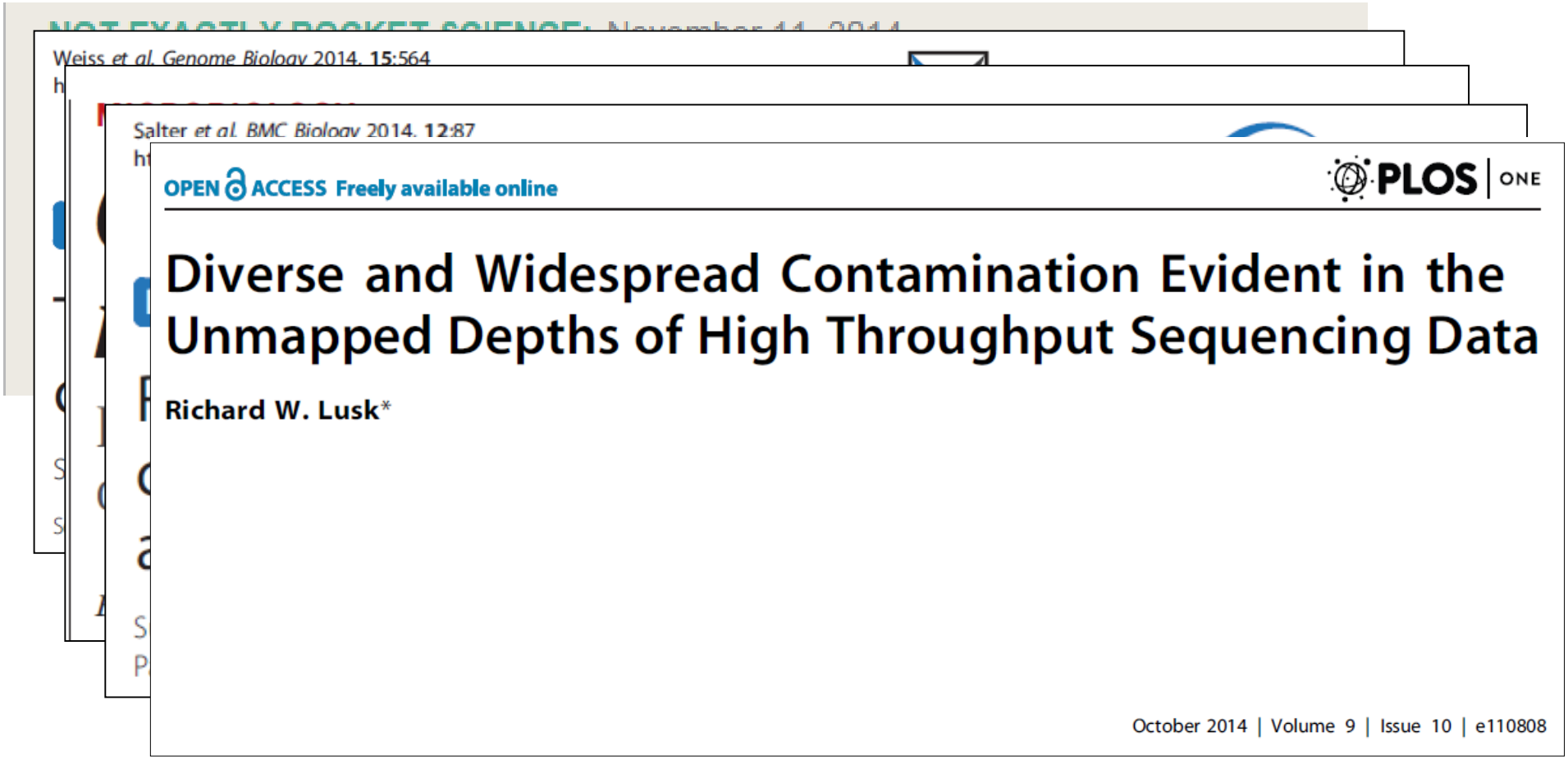
Plate42



Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate



Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Table 1 List of contaminant genera detected in sequenced negative 'blank' controls

Phylum	List of constituent contaminant genera
Proteobacteria	Alpha-proteobacteria:
	Afipia, Aquabacterium[e], *Asticcacaulis, Aurantimonas, Beijerinckia, Bosea, Bradyrhizobium*[d], *Brevundimonas*[c], *Caulobacter, Craurococcus, Devosia, Hoeflea*[e], *Mesorhizobium, Methylobacterium*[c], *Novosphingobium, Ochrobactrum, Paracoccus, Pedomicrobium, Phyllobacterium*[e], *Rhizobium*[c,d], *Roseomonas, Sphingobium, Sphingomonas*[c,d,e], *Sphingopyxis*
	Beta-proteobacteria:
	Acidovorax[c,e], *Azoarcus*[e], *Azospira, Burkholderia*[d], *Comamonas*[c], *Cupriavidus*[c], *Curvibacter, Delftia*[e], *Duganella*[a], *Herbaspirillum*[a,c], *Janthinobacterium*[e], *Kingella, Leptothrix*[a], *Limnobacter*[e], *Massilia*[c], *Methylophilus, Methyloversatilis*[e], *Oxalobacter, Pelomonas, Polaromonas*[e], *Ralstonia*[b,c,d,e], *Schlegelella, Sulfuritalea, Undibacterium*[e], *Variovorax*
	Gamma-proteobacteria:
	Acinetobacter[a,d,c], *Enhydrobacter, Enterobacter, Escherichia*[a,c,d,e], *Nevskia*[e], *Pseudomonas*[b,d,e], *Pseudoxanthomonas, Psychrobacter, Stenotrophomonas*[a,b,c,d,e], *Xanthomonas*[b]
Actinobacteria	*Aeromicrobium, Arthrobacter, Beutenbergia, Brevibacterium, Corynebacterium, Curtobacterium, Dietzia, Geodermatophilus, Janibacter, Kocuria, Microbacterium, Micrococcus, Microlunatus, Patulibacter, Propionibacterium*[e], *Rhodococcus, Tsukamurella*
Firmicutes	*Abiotrophia, Bacillus*[b], *Brevibacillus, Brochothrix, Facklamia, Paenibacillus, Streptococcus*
Bacteroidetes	*Chryseobacterium, Dyadobacter, Flavobacterium*[d], *Hydrotalea, Niastella, Olivibacter, Pedobacter, Wautersiella*
Deinococcus-Thermus	*Deinococcus*
Acidobacteria	Predominantly unclassified Acidobacteria Gp2 organisms

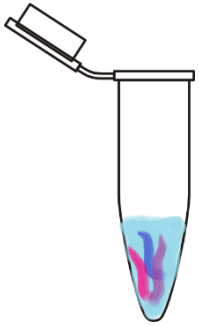
The listed genera were all detected in sequenced negative controls that were processed alongside human-derived samples in our laboratories (WTSI, ICL and UB) over a period of four years. A variety of DNA extraction and PCR kits were used over this period, although DNA was primarily extracted using the FastDNA SPIN Kit for Soil. Genus names followed by a superscript letter indicate those that have also been independently reported as contaminants previously. [a]also reported by Tanner *et al.* [12]; [b]also reported by Grahn *et al.* [14]; [c]also reported by Barton *et al.* [17]; [d]also reported by Laurence *et al.* [18]; [e]also detected as contaminants of multiple displacement amplification kits (information provided by Paul Scott, Wellcome Trust Sanger Institute). ICL, Imperial College London; UB, University of Birmingham; WTSI, Wellcome Trust Sanger Institute.

Salter et al., (2014), *BMC Biology*; **12**; 1-12

Steps involved:

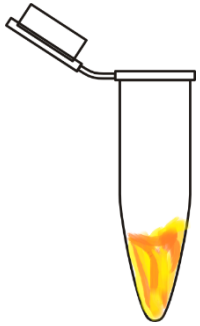
Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate



Potential contaminating genera from PCR reagents

+



Biological sample DNA

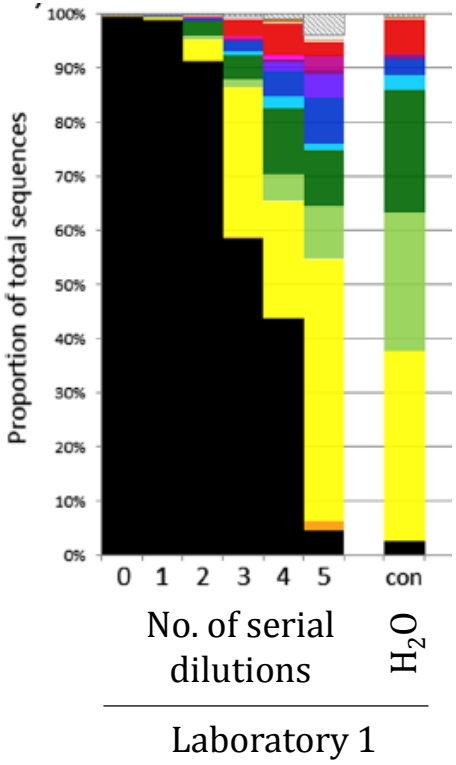
=

Background noise present in biological sample sequencing results

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate



Targeted the 16S rRNA gene of *Salmonella bongori* (in black) which had undergone five rounds of serial ten-fold dilutions

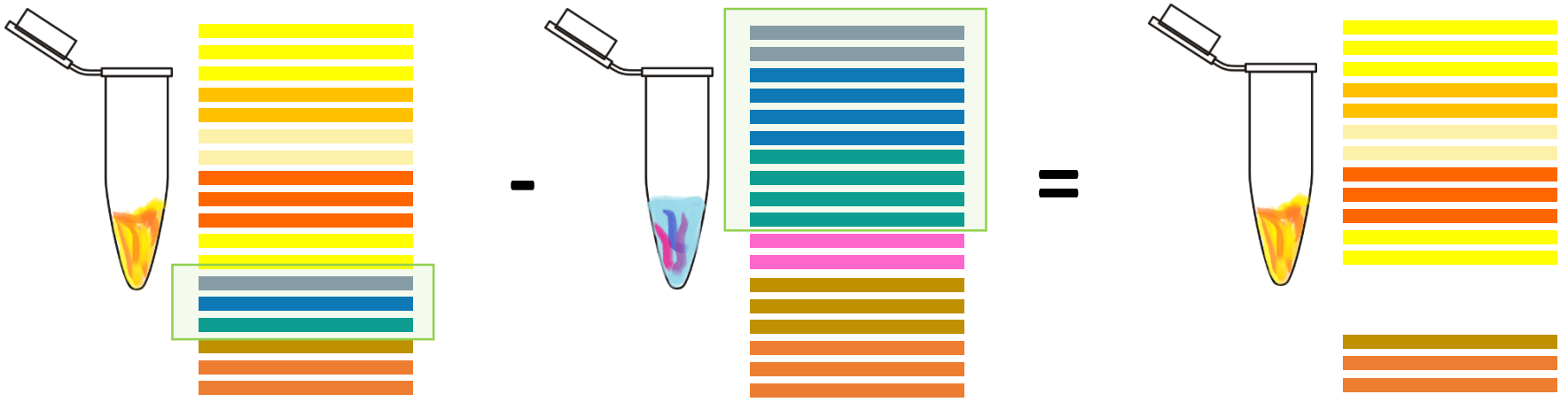
Salter et al., (2014), *BMC Biology*; **12**; 1-12

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Clearly we need to be removing "background noise" added to our reactions via reagents such as PCR grade water.
However, removing the reads generated from no-template controls themselves will result in over-compensating for contamination for example:



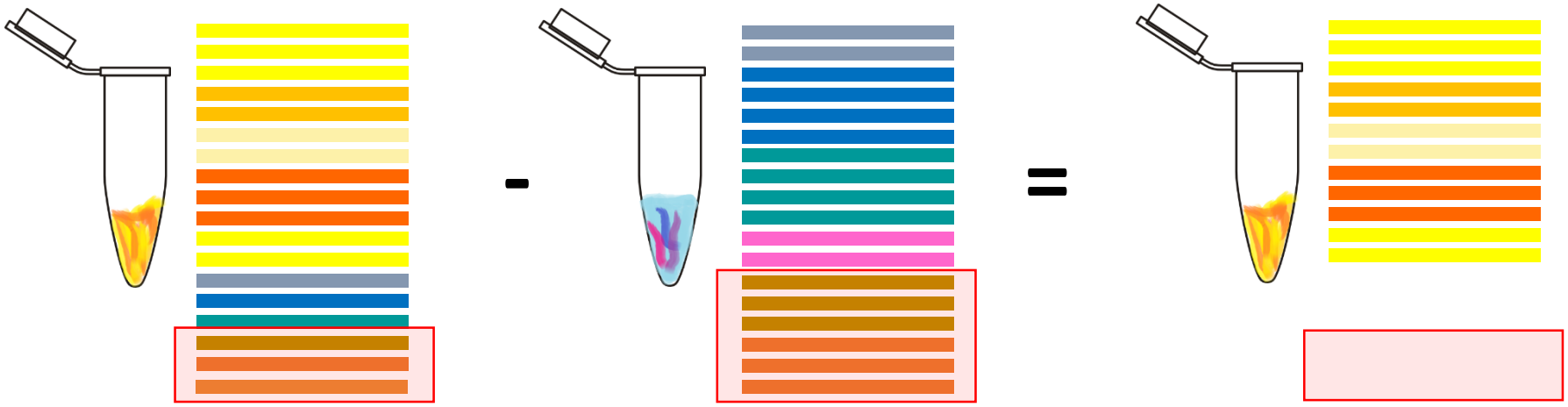
Biological sample - No-template control = Biological sample

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Clearly we need to be removing "background noise" added to our reactions via reagents such as PCR grade water.
However, removing the reads generated from no-template controls themselves will result in over-compensating for contamination for example:



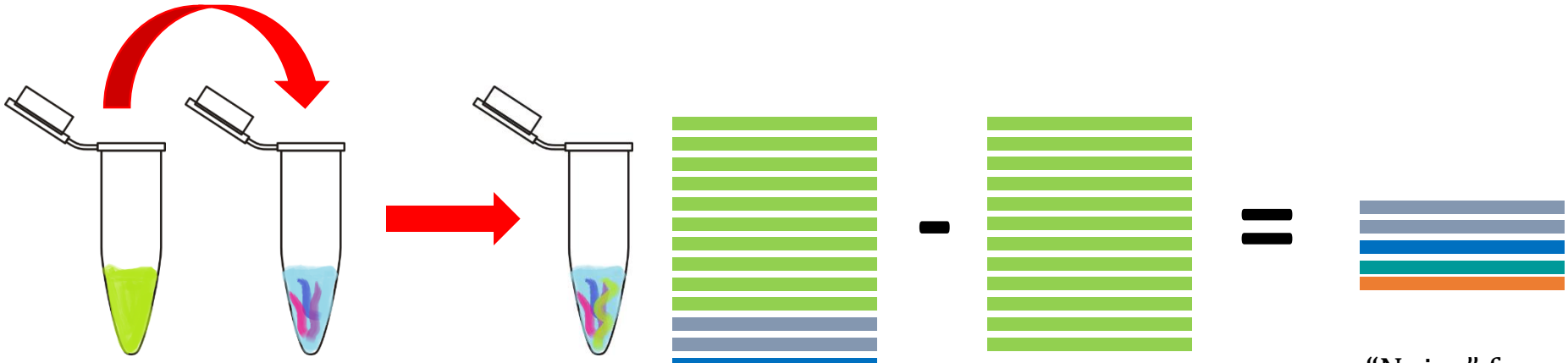
Biological sample | No-template control | Biological sample

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

We can control for this by spiking our no-template controls with "competing" DNA of a known organism
The spike-DNA should be from an organism not expected in the no-template control
The spike-DNA should be at a **concentration representative of that of the biological samples**



Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Biological sample

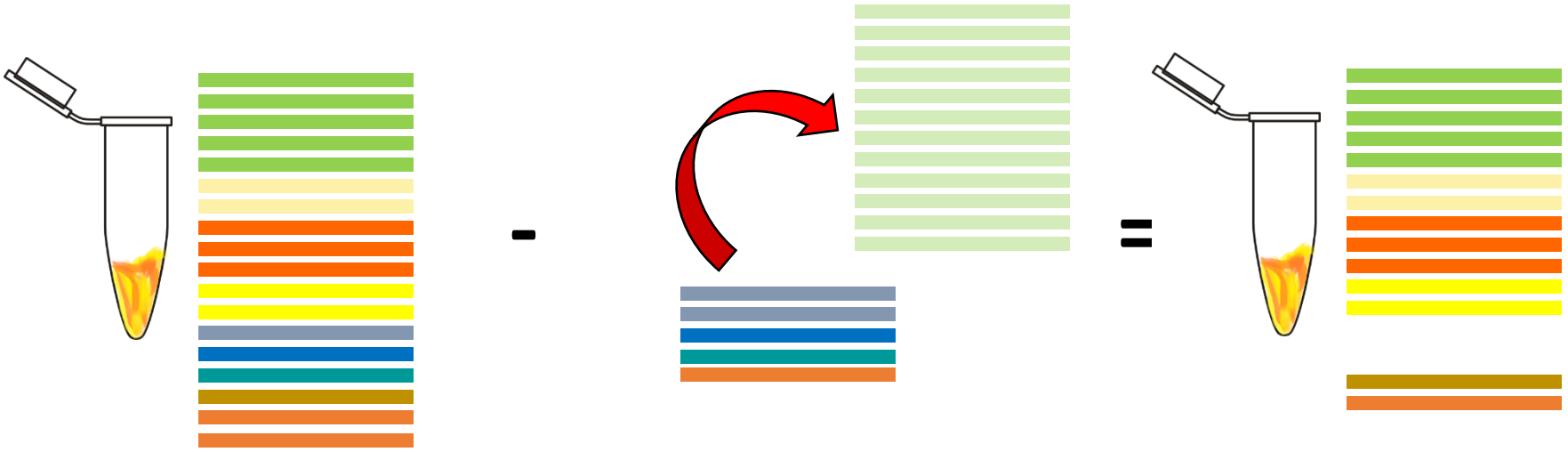
“Noise” following removal of spike-in DNA

Biological sample without over-compensating for contamination

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate



Biological sample

“Noise” following removal of spike-in DNA

Biological sample without over-compensating for contamination

Steps involved:

Step 1
Design your experiment:
DNA extraction protocol

Quantify template using 16S rRNA qPCR

The 16S rRNA real time PCR protocol is published at:
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0017035

301460	F05	Unkn		24.26819	0.116
Primestore_FTDint_P1 (SC 9 B8)	F06	Unkn	Primestore_int	36.03133	0.000
502527	F07	Unkn		29.54712	0.003
101398	F08	Unkn		29.55681	0.003
102598	F09	Unkn		23.255	0.233
104454	F10	Unkn		24.64554	0.089
518306	F11	Unkn		22.84429	0.310
302194	F12	Unkn		21.13398	1.015
102539	G03	Unkn		33.36721	0.000
104435	G04	Unkn		21.43081	0.910
118685	G03	Unkn		24.20348	0.121
104212	G04	Unkn		19.27805	3.674
108148	G05	Unkn		24.50339	0.098
w ater	G06	Unkn	MilliQ	35.66736	0.000
505016	G07	Unkn		35.65833	0.000
103051	G08	Unkn		27.24958	0.015
104614	G09	Unkn		19.42097	3.327
108290	G10	Unkn		29.70401	0.003
512561	G11	Unkn		24.04204	0.135
303017	G12	Unkn		20.00377	2.222
507636	H03	Unkn		21.52387	0.854
508423	H04	Unkn		25.33993	0.066
506335	water	Unkn	MilliQ	36.39297	0.000
301009	H04	Unkn		21.97896	0.565

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

Plate 1

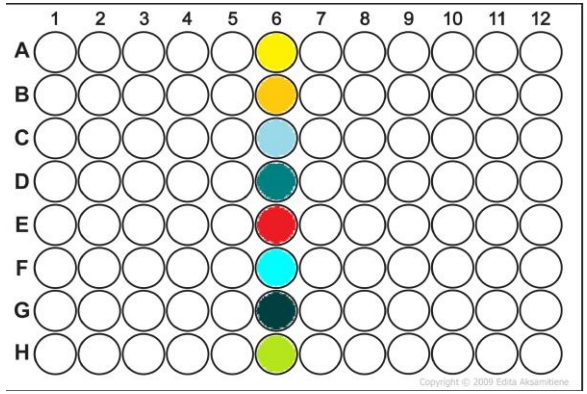


Plate 2

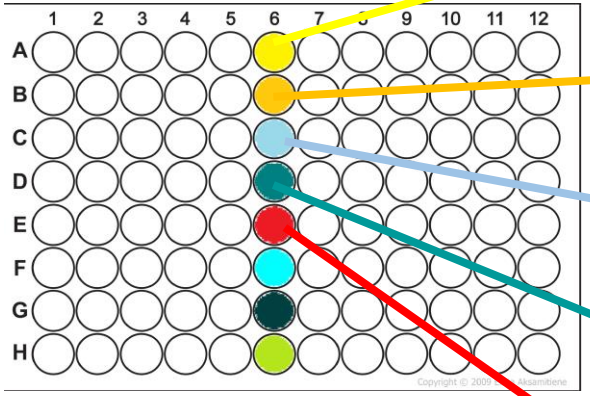


Plate 3

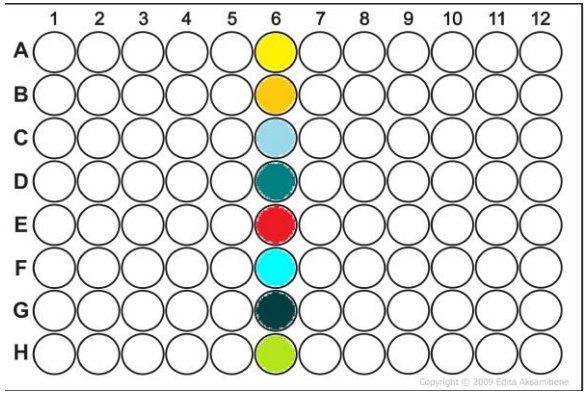
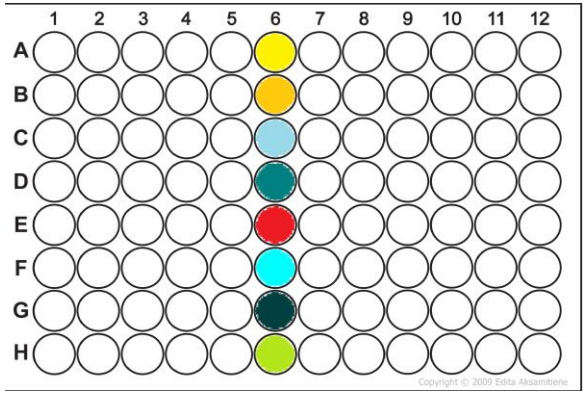


Plate42



Between-run repeat (process any specimen in duplicate per run to measure reproducibility across runs)

Within-run repeat (process any specimen in duplicate on each of the plates to measure reproducibility)

Water used during PCR (to determine if any contamination was introduced during PCR reactions)

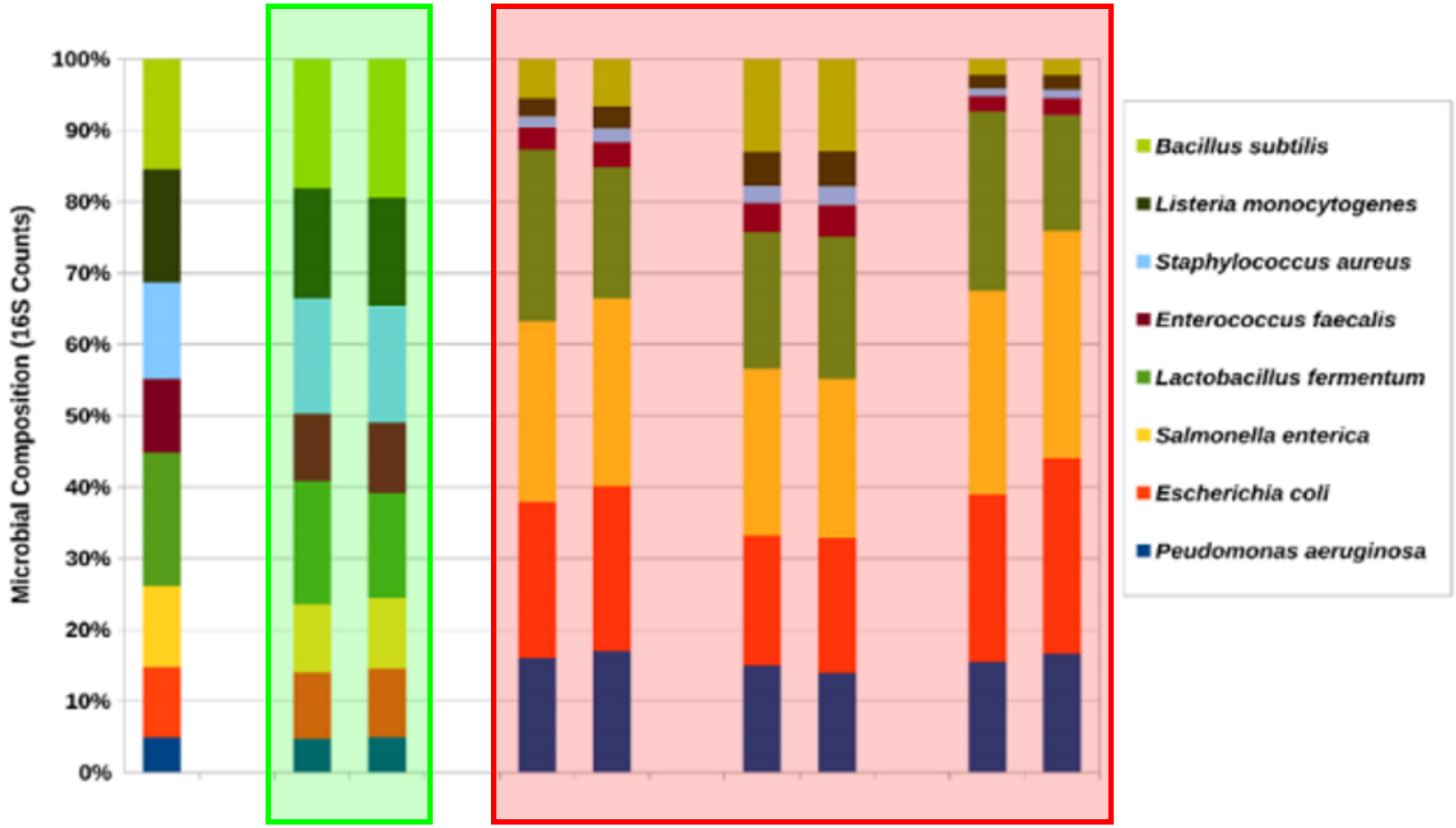
Water spiked with known bacterial DNA (to control for contaminants introduced during PCR reactions)

Bacterial mock community DNA (mix of bacterial DNA serves as sequencing control)

Steps involved:

Step 1
Design your experiment:
Include controls

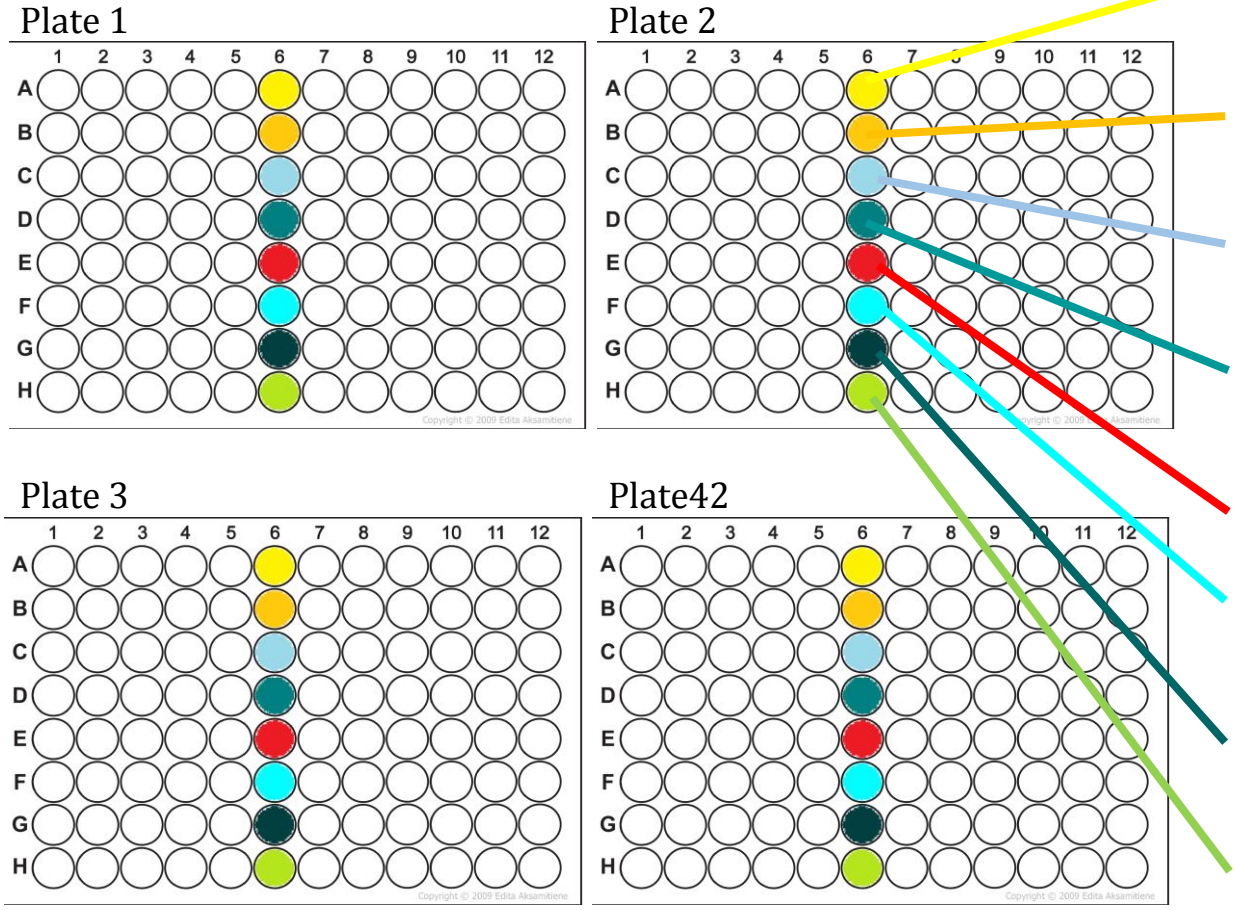
Include a set of controls on each 96-well plate



Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate



Between-run repeat (process any specimen in duplicate per run to measure reproducibility across runs)

Within-run repeat (process any specimen in duplicate on each of the plates to measure reproducibility)

Water used during PCR (to determine if any contamination was introduced during PCR reactions)

Water spiked with known bacterial DNA (to control for contaminants introduced during PCR reactions)

Bacterial mock community DNA (mix of bacterial DNA serves as sequencing control)

Sample storage medium or elution buffer (to determine if contaminants was introduced during sample collection)

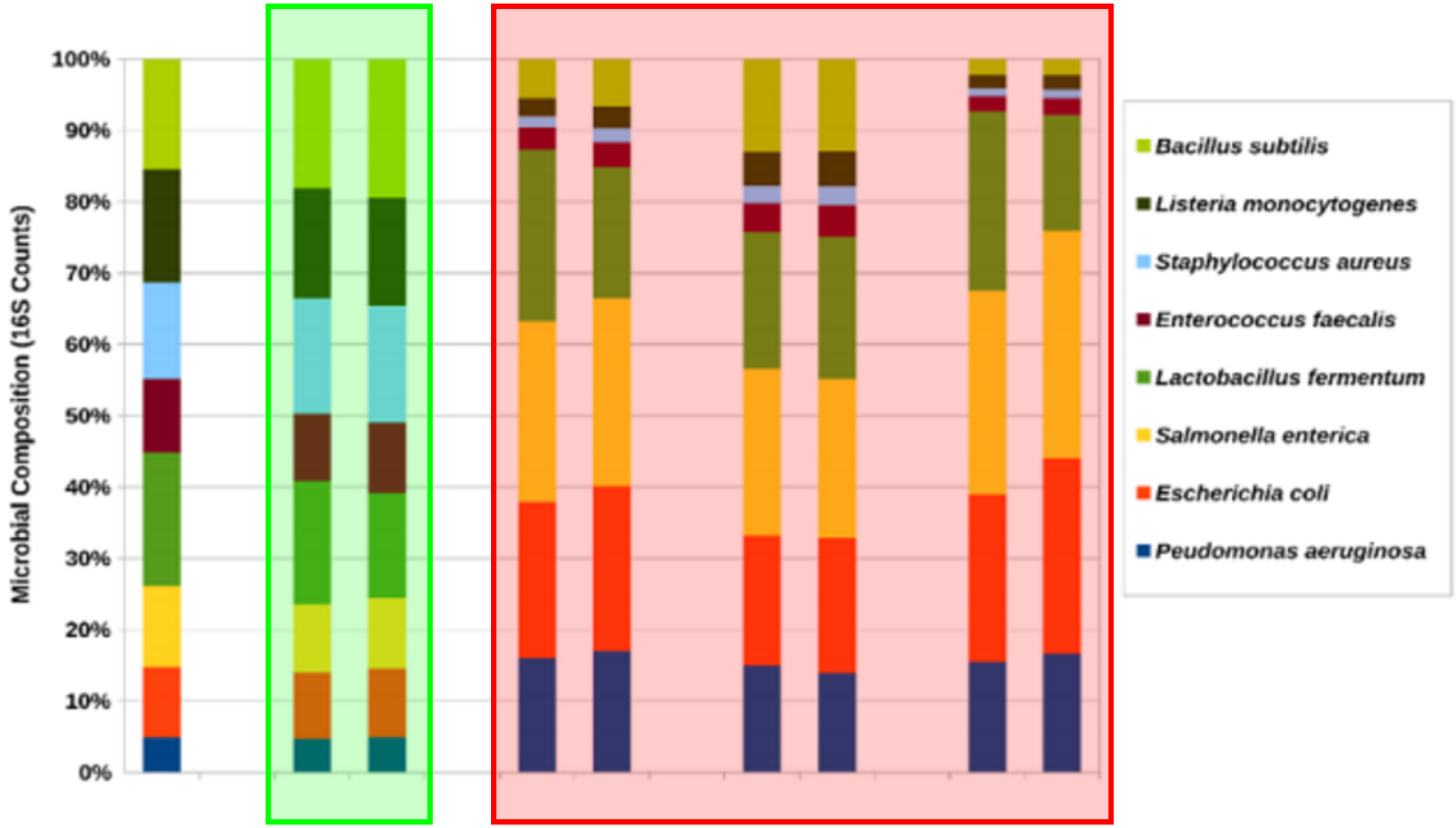
Sample storage medium or elution buffer spiked (to control for contaminants introduced during sample collection)

Extraction control (to determine the extraction efficiency)

Steps involved:

Step 1
Design your experiment:
Include controls

Include a set of controls on each 96-well plate

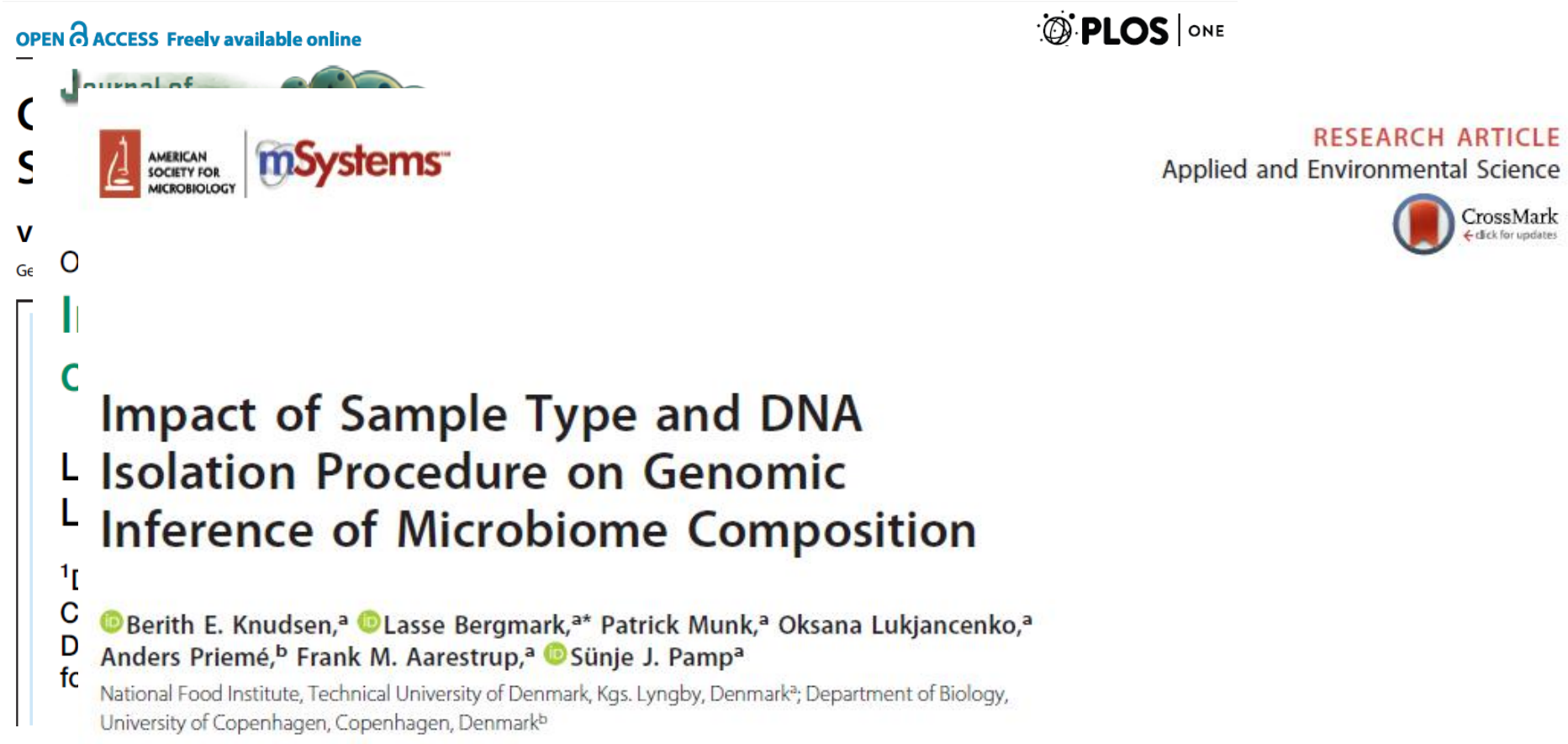


Steps involved:

Step 1
Design your experiment:
DNA extraction protocol

Use an optimised DNA extraction protocol for your sample type

Different DNA extraction protocols will produce differences in microbiota communities:



OPEN ACCESS Freely available online

PLOS ONE

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems

RESEARCH ARTICLE
Applied and Environmental Science

CrossMark click for updates

Impact of Sample Type and DNA Isolation Procedure on Genomic Inference of Microbiome Composition

Berith E. Knudsen,[a] Lasse Bergmark,[a]* Patrick Munk,[a] Oksana Lukjancenko,[a] Anders Priemé,[b] Frank M. Aarestrup,[a] Sünje J. Pamp[a]

National Food Institute, Technical University of Denmark, Kgs. Lyngby, Denmark[a]; Department of Biology, University of Copenhagen, Copenhagen, Denmark[b]

Steps involved:

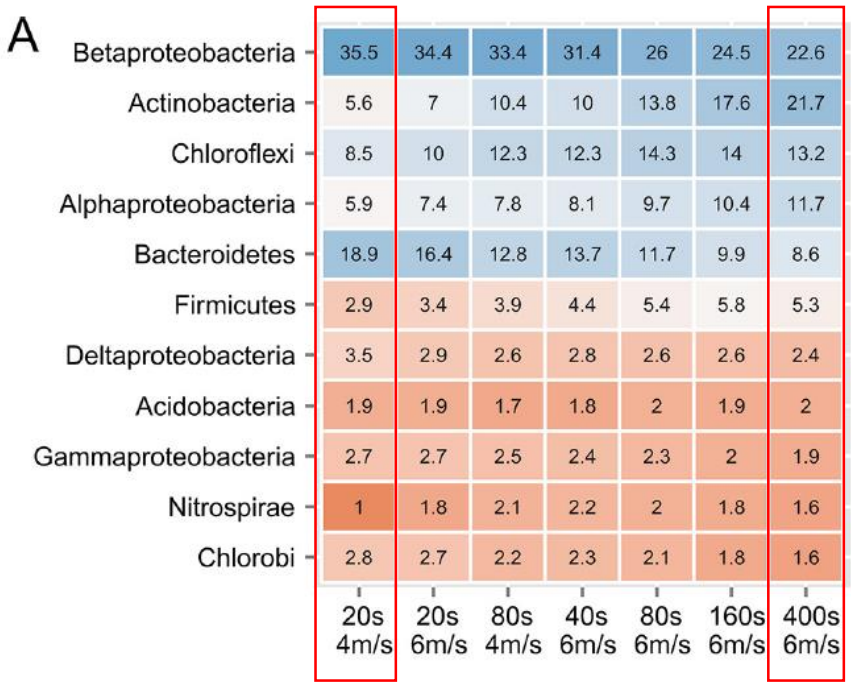
Step 1
Design your experiment:
DNA extraction protocol

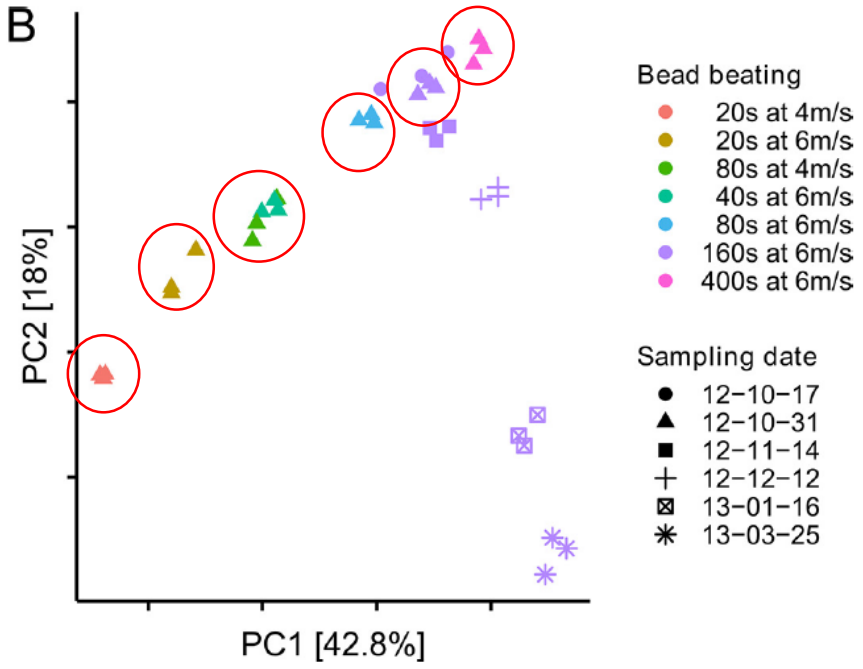
Use an optimised DNA extraction protocol for your sample type

Another example is the effect of mechanical lysis methods:

Increased bead beating had a dramatic influence on the observed community composition

Compared to the time series samples, the effect of bead beating was larger than the effect of sampling 5 months apart.





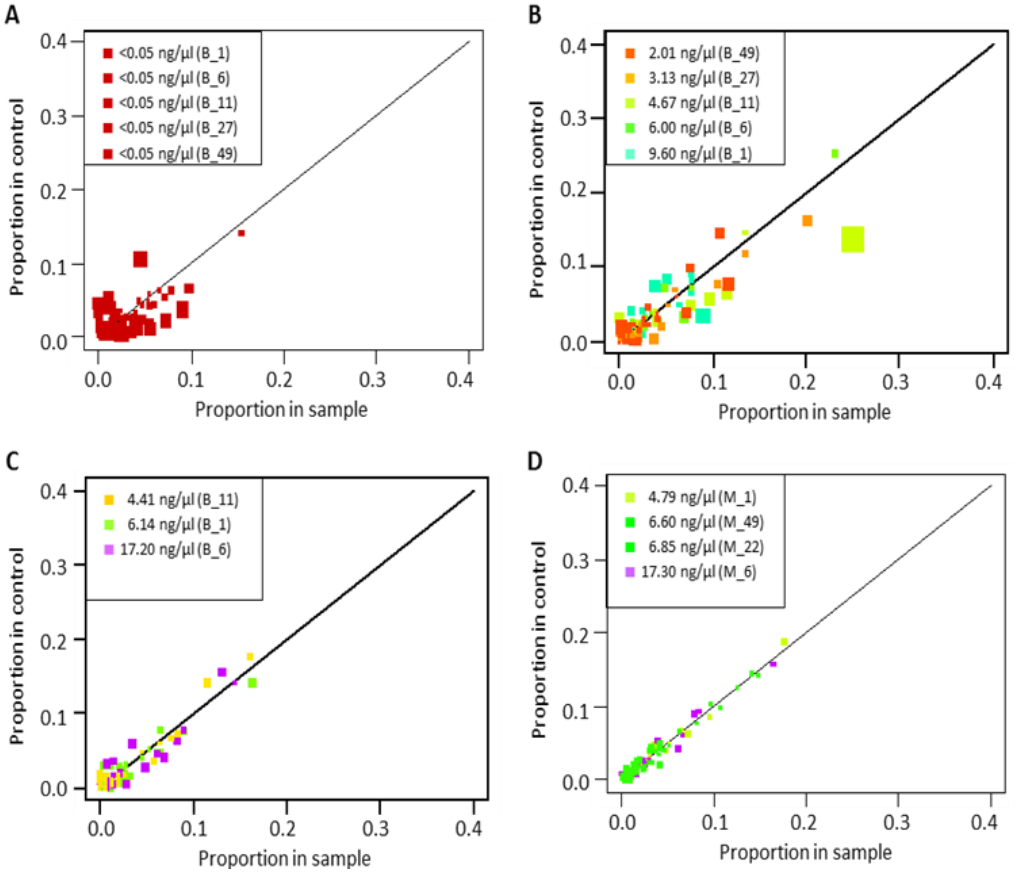
Albertsen et al., (2015), *PLoS ONE* 10(7): e0132783

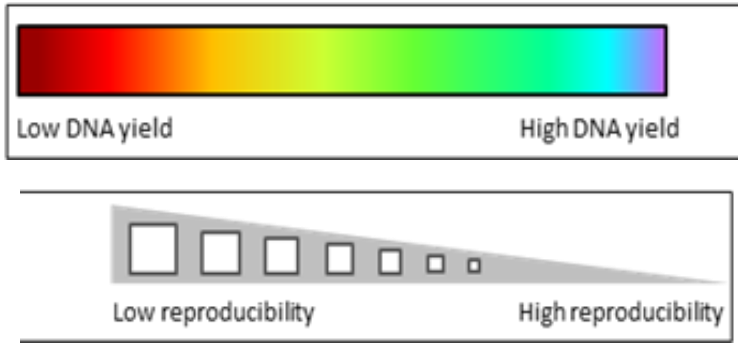
Steps involved:

Step 1
Design your experiment:
DNA extraction protocol

Use an optimised DNA extraction protocol for your sample type

DNA concentrations also seem to influence sequencing reproducibility:





Steps involved:

Step 1
Design your experiment:
DNA extraction protocol

Batch extractions based on your sequencing plates layout

Plate 1 sequencing layout

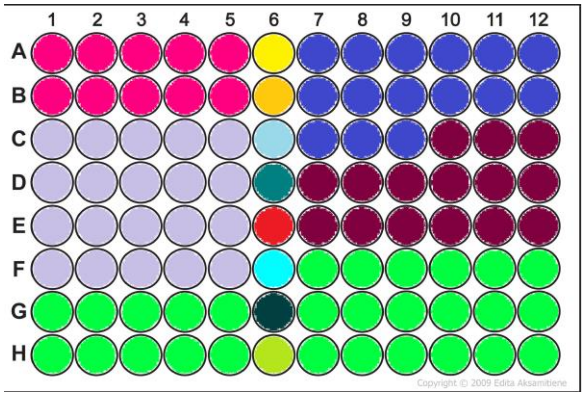


Plate 1 extraction layout

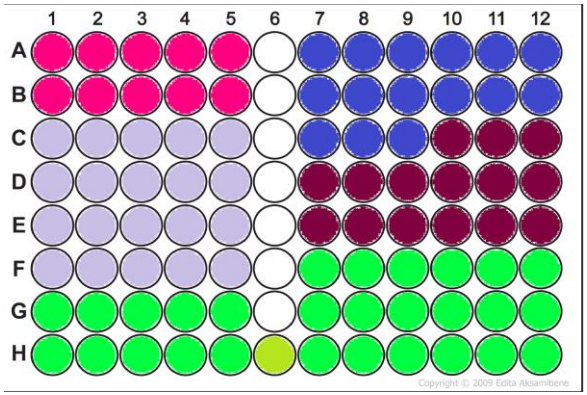


Plate 1 sequencing layout

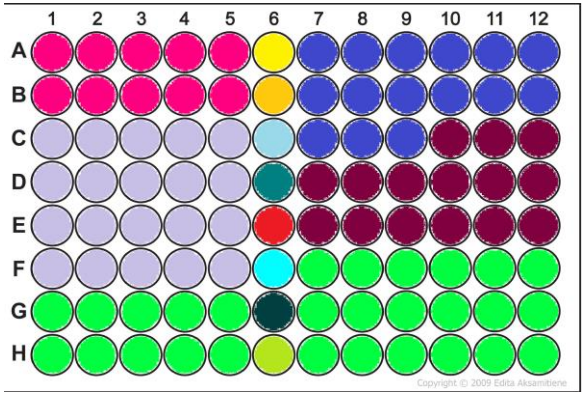
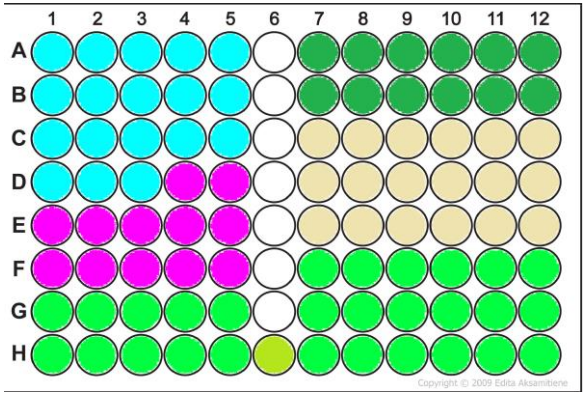


Plate 1 extraction layout

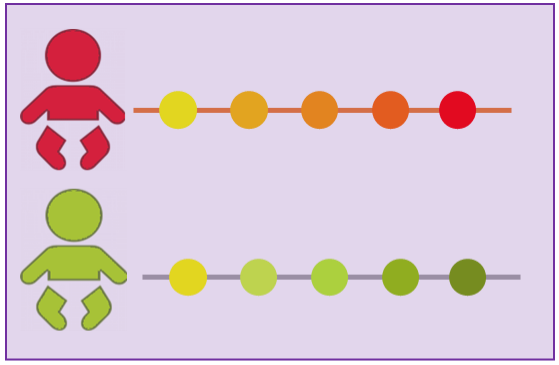


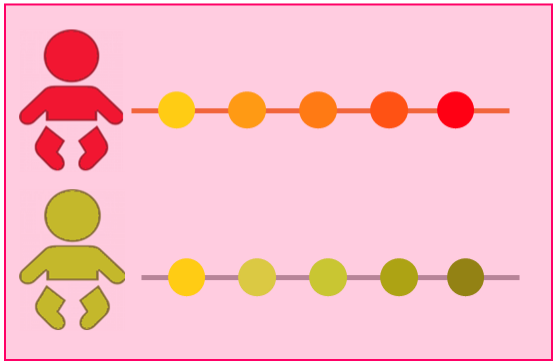
Steps involved:

Step 1
Design your experiment:
based on your research question

Think about your study design

Case-control set 1





Case-control set 2

Plate 1

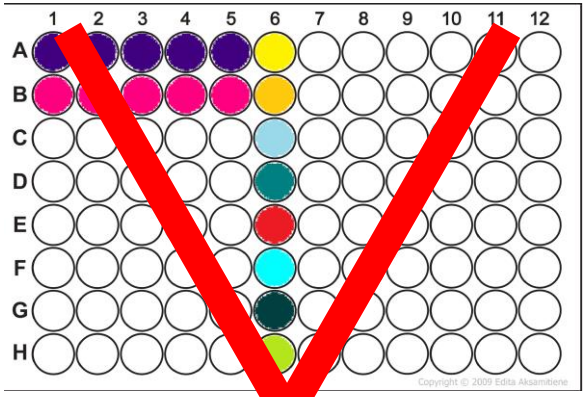


Plate 2

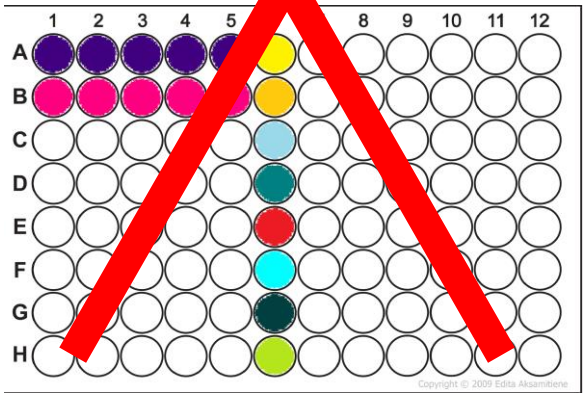


Plate 1

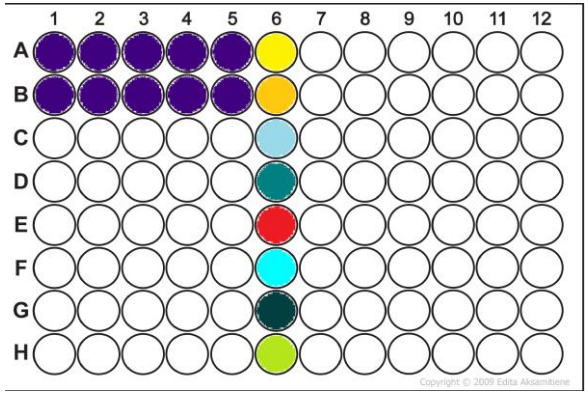
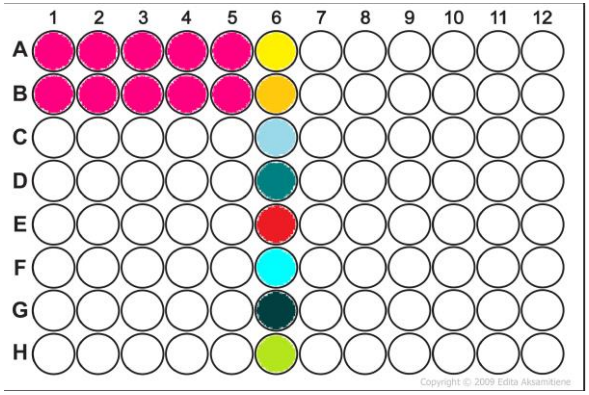


Plate 2



Steps involved:

Step 2
PCR amplification

Selection of primers also influence microbial profiles sequenced

DOI 10.1186/s40168-015-0087-4

Microbiome

RESEARCH Open Access

16S rRNA gene-based profiling of the human infant gut microbiota is strongly influenced by sample processing and PCR primer choice

Alan W. Walker[1,2], Jennifer C. Martin[1], Paul Scott[2], Julian Parkhill[2], Harry J. Flint[1] and Karen P. Scott[1*]

Abstract

Background: Characterisation of the bacterial composition of the gut microbiota is increasingly carried out with a view to establish the role of different bacterial species in causation or prevention of disease. It is thus essential that the methods used to determine the microbial composition are robust. Here, several widely used molecular techniques were compared to establish the optimal methods to assess the bacterial composition in faecal samples from babies, before weaning.

Results: The bacterial community profile detected in the faeces of infants is highly dependent on the methodology used. Bifidobacteria were the most abundant bacteria detected at 6 weeks in faeces from two initially breast-fed babies using fluorescent in situ hybridisation (FISH), in agreement with data from previous culture-based studies. Using the 16S rRNA gene sequencing approach, however, we found that the detection of bifidobacteria in particular crucially depended on the optimisation of the DNA extraction method, and the choice of primers used to amplify the V1–V3 regions of 16S rRNA genes prior to subsequent sequence analysis. Bifidobacteria were only well represented among amplified 16S rRNA gene sequences when mechanical disruption (bead-beating) procedures for DNA extraction were employed together with optimised "universal" PCR primers. These primers incorporate degenerate bases at positions where mismatches to bifidobacteria and other bacterial taxa occur.

Steps involved:

Step 2
PCR amplification

We perform a 2-step PCR to amplify the V4 region of the 16S rRNA gene

Short PCR (10 cycles)

We use 4ul of template for stool samples and 7ul of template for nasopharyngeal and breast milk samples

515F: GTGCCAGCHGCYGCGGT
806R: GGACTACNNGGGTWTCTAAT

Primers are modified primers from: Caporaso et al., *PNAS.*, 2011; 108:4516-4522

Long PCR

We use 4ul of amplicon from the short PCR for stool samples (and 7ul for nasopharyngeal and breast milk samples) as template in the long PCR

515Fmod4_std_12-16N

5' AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNNNNNNNN(12-16)GTGCCAGCHGCYGCGGT 3' →

Sequence adapter | Sequencing priming region | Random nucleotides | 16S primer

806Rmod1_std_12-16N_BC001

← 3' TAATCTWTGGGNNCATCAGG(12-16)NNNNNNNNNNNNTCTAGCCTTCTCGTGTGCACAGTTGAGGTCAGTGTTAGTCAGAGCATAGAGCATAGAGCATACGGCAGAAGACGAAC 5'

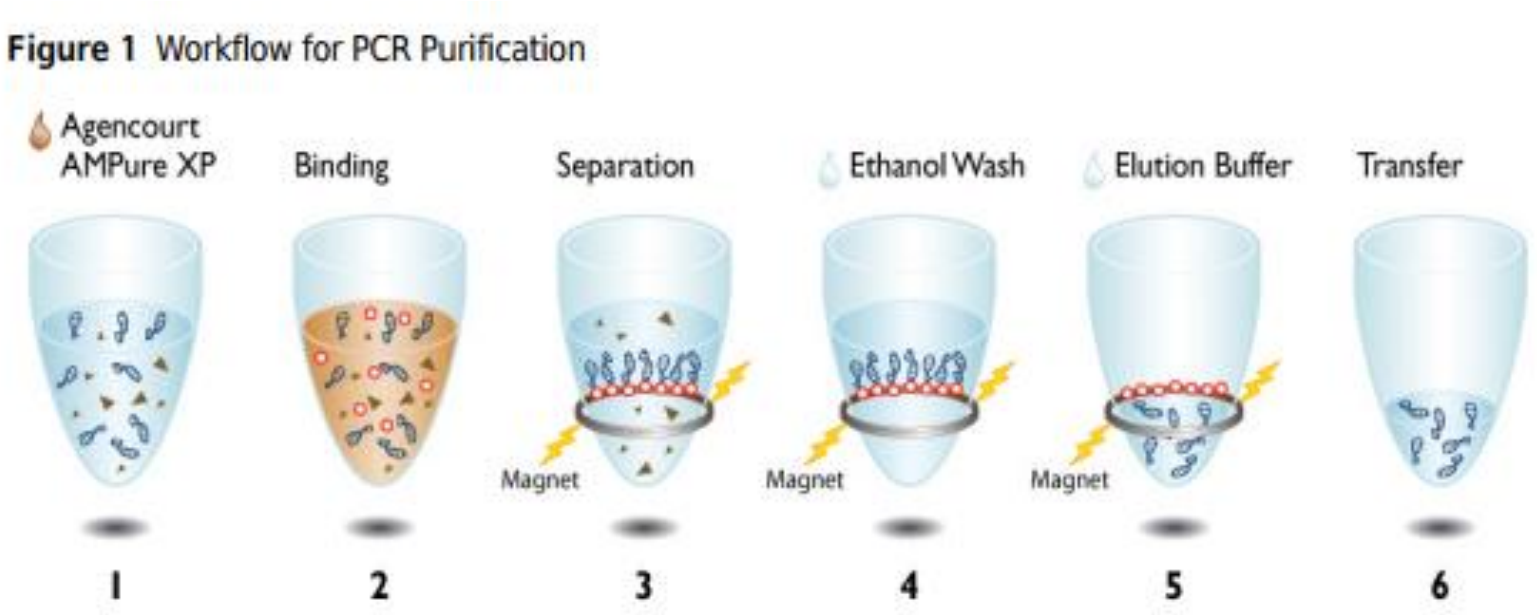
16s primer | Random nucleotides | Sequencing priming region | Barcode | Sequence adapter

Steps involved:

Step 3
Cleaning of amplicon products

The Agencourt AMPure XP PCR Purification System

Agencourt AMPure XP utilizes an optimized buffer to selectively bind DNA fragments 100 bp and larger to paramagnetic beads.
Excess primers, nucleotides, salts, and enzymes can be removed using a simple washing procedure.
The result is a more purified PCR product.



Steps involved:

Step 4
Gel electrophoresis and quantification of cleaned amplicon product

This step is used to assess which samples and what volumes will be pooled

Plate 1

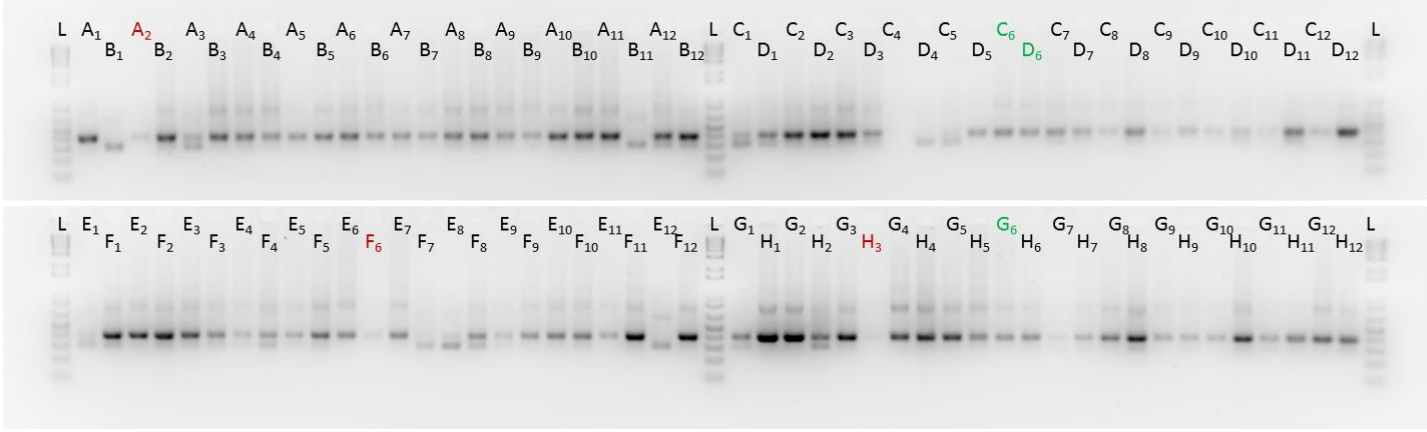
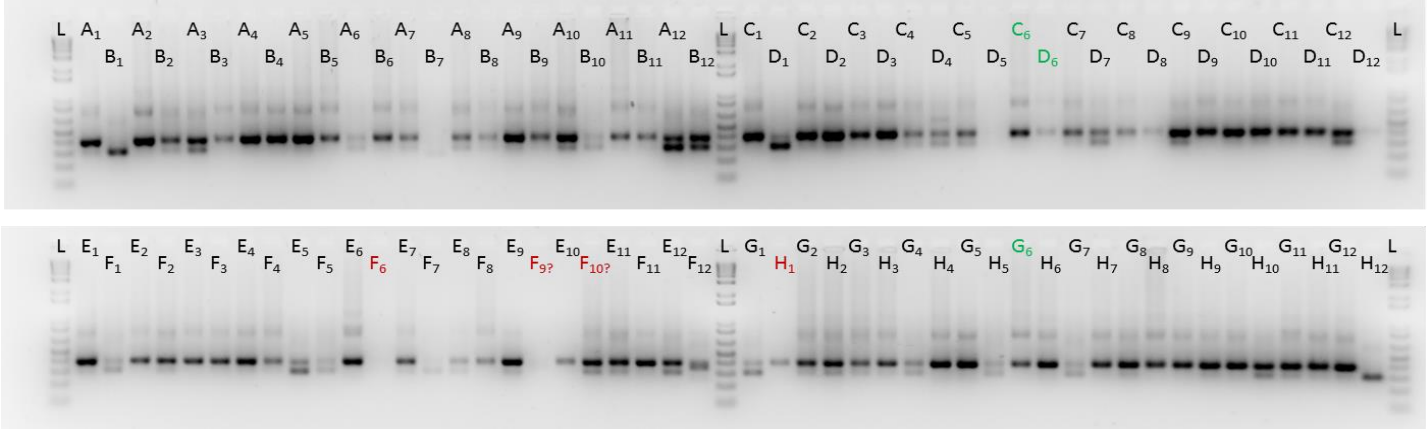


Plate 2



position	gel_score
P1_A01	positive
P1_A02	faint
P1_A03	positive
P1_A04	positive
P1_A05	positive
P1_A06	positive
P1_A07	positive
P1_A08	positive
P1_A09	positive
P1_A10	positive
P1_A11	positive
P1_A12	positive
P1_B01	positive
P1_B02	positive
P1_B03	positive
P1_B04	positive
P1_B05	positive
P1_B06	positive
P1_B07	positive
P1_B08	positive
P1_B09	faint
P1_B10	positive
P1_B11	negative
P1_B12	positive

Steps involved:

Step 4
Gel electrophoresis and quantification of cleaned amplicon product

This step is used to assess which samples and what volumes will be pooled

Plate 1	1	2	3	4	5	6	7	8	9	10	11	12
A	1.23	9.43	5.32	2.45	6.77	4.07	6.18	4.00	4.71	1.76	2.49	1.87
B	4.10	2.77	2.32	4.33	2.75	6.02	4.51	3.08	4.96	2.08	5.54	1.69
C	4.31	1.81	2.14	8.33	7.91	5.48	4.35	2.67	6.37	2.86	3.40	1.72
D	3.68	3.06	4.33	6.39	3.27	5.31	7.01	1.95	7.37	6.56	2.35	2.45
E	5.78	1.34	2.93	4.84	2.58	4.00	4.02	3.34	7.99	4.76	2.58	4.77
F	1.36	2.30	3.71	4.40	1.82	37.55	9.65	4.87	7.01	3.23	2.49	1.52
G	4.25	1.50	2.48	1.90	2.20	5.88	31.28	2.70	6.60	5.78	1.65	1.33
H	1.18	2.35	63.67	1.72	3.93	7.11	4.23	1.89	3.52	1.82	2.55	1.57

Plate 2	1	2	3	4	5	6	7	8	9	10	11	12
A	1.27	1.02	1.68	1.51	1.23	9.78	5.70	4.02	2.47	2.25	1.85	2.37
B	1.81	2.43	2.36	1.82	1.50	2.55	109.21	7.64	4.06	6.29	2.45	1.46
C	1.18	1.36	2.40	5.93	4.31	2.96	3.66	9.45	2.15	2.30	3.32	1.95
D	1.83	1.16	1.57	6.41	14.95	5.28	4.90	11.21	2.96	2.37	1.61	21.58
E	0.99	1.42	1.93	1.83	3.77	2.27	2.78	17.45	2.72	9.26	1.93	2.53
F	3.46	1.70	1.77	4.81	5.01	-191.47	10.34	4.92	-669.00	6.45	1.52	3.86
G	3.70	2.11	2.68	5.84	2.07	3.01	7.84	3.15	2.95	1.64	1.47	1.89
H	7.17	1.48	1.68	1.39	5.64	1.79	2.91	2.38	1.68	1.88	1.59	2.36

Plate 3	1	2	3	4	5	6	7	8	9	10	11	12
A	2.51	0.55	0.80	0.95	0.81	6.37	0.57	0.58	0.70	1.13	75.70	2.47
B	0.63	0.59	0.54	0.94	0.73	1.07	0.73	0.84	0.73	0.69	1.92	1.45
C	4.40	0.71	0.80	0.77	0.81	0.64	1.97	1.32	0.84	1.06	1.94	2.45
D	0.75	1.14	0.74	0.70	0.88	0.73	1.10	1.24	1.40	1.19	2.29	1.72
E	2.95	1.00	0.61	3.08	3.19	0.74	1.11	3.06	1.12	1.15	2.79	2.20
F	0.97	0.55	0.58	0.64	0.73	3.48	26.25	2.73	1.37	1.49	1.57	1.34
G	0.96	0.64	0.96	0.64	0.96	0.67	3.37	1.09	0.90	1.91	1.73	2.24
H	2.35	1.10	0.68	0.60	0.57	0.76	2.09	1.21	2.32	1.92	1.13	1.73

Plate 4	1	2	3	4	5	6	7	8	9	10	11	12
A	1.09	1.10	1.93	1.87	1.80	2.15	1.69	1.73	2.19	2.69	1.54	1.42
B	3.40	3.13	1.54	1.25	1.37	2.72	3.86	2.30	6.13	2.26	2.77	2.78
C	1.38	1.75	1.37	1.37	1.83	1.61	1.75	3.01	2.20	2.30	2.77	2.52
D	1.90	4.34	1.20	4.37	2.03	1.33	1.57	2.36	1.51	2.76	3.18	1.96
E	2.90	2.29	1.66	1.40	1.89	2.08	2.88	2.92	1.93	1.75	3.32	2.20
F	1.65	1.43	1.84	1.97	2.66	45.62	28.16	2.69	2.89	2.19	3.99	2.15
G	1.34	1.92	1.66	1.88	1.72	1.69	2.05	1.84	3.29	2.35	2.06	2.09
H	1.58	1.96	1.43	1.32	1.83	1.43	3.68	2.20	2.14	2.23	2.52	2.21

Double stranded DNA is quantified using the QuantiFluor™ dsDNA System

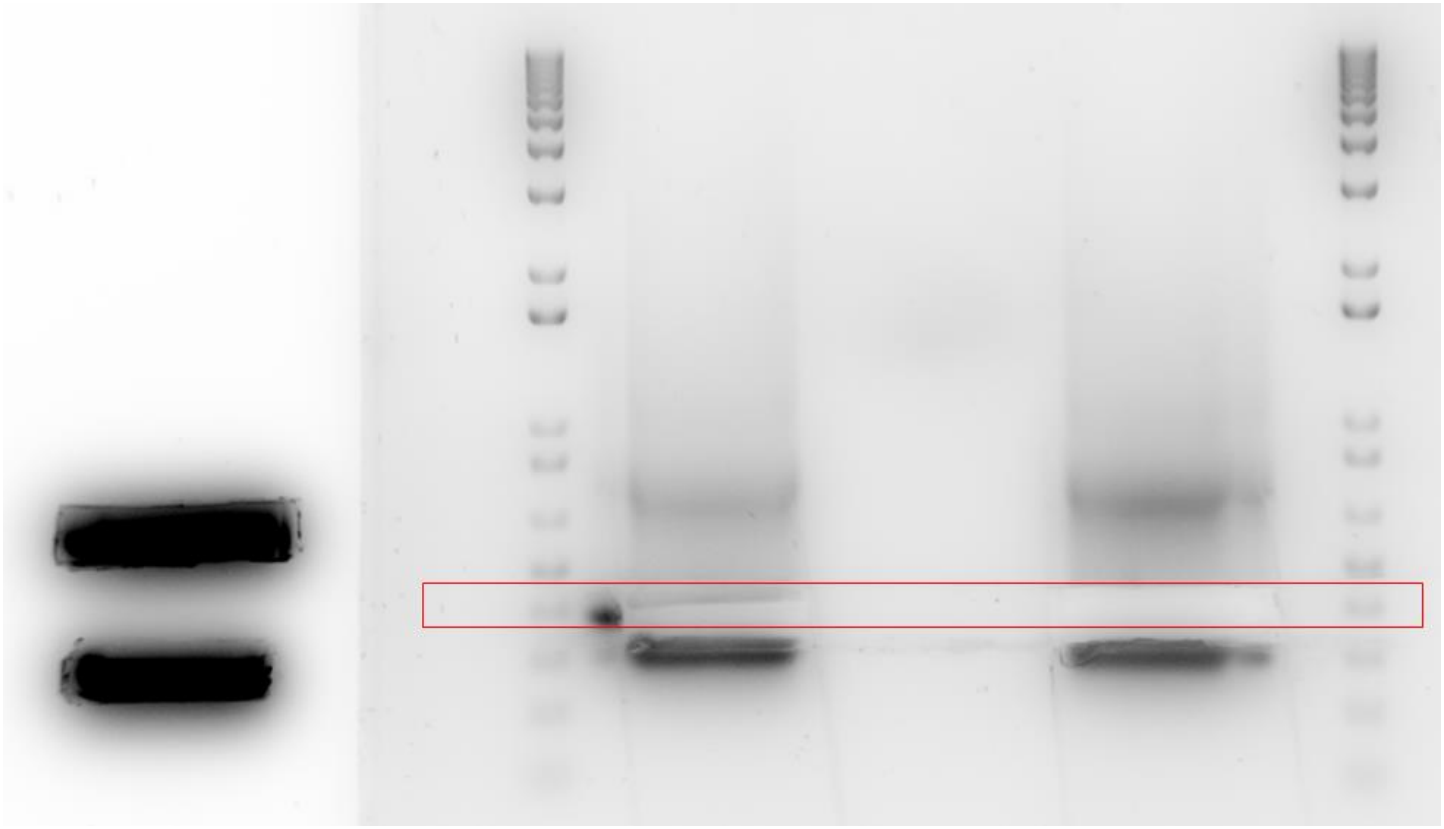
Equimolar amounts of each amplicon are pooled to achieve appropriate numbers of sequences per sample

At this stage any samples failing the previous QC steps are either excluded from the pool or flagged for downstream assessment

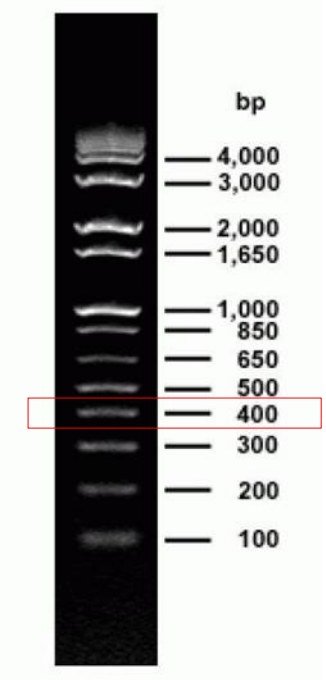
Steps involved:

Step 5
Agarose gel extraction of pooled product

This step is used to extract the pooled amplicon library of interest for sequencing (removal of primer dimers and non-specific binding)



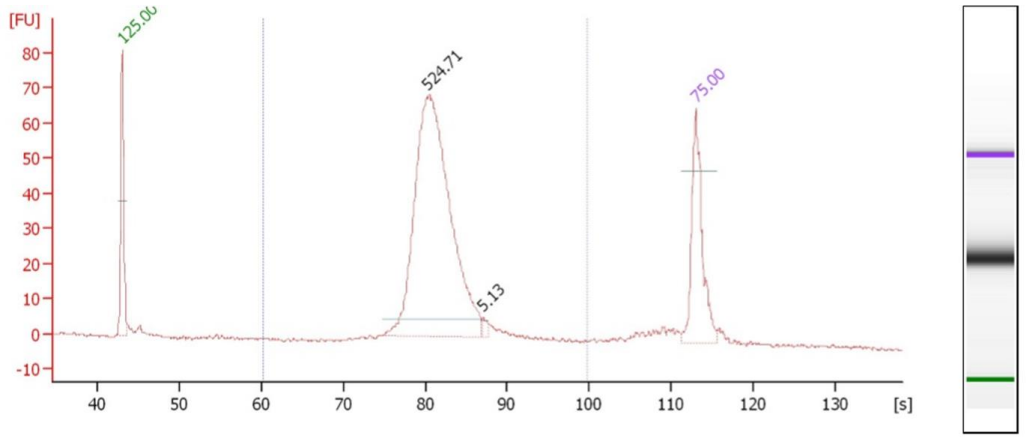
TrackIt™ 1 Kb Plus DNA Ladder



Steps involved:

Step 6
Cleaning of extracted amplicon library and final QC steps

- ***The amplicon product is cleaned using the QIAquick Gel Extraction Kit (Qiagen) to remove agarose gel and impurities.***
- ***Quantification of the cleaned extracted library is then performed using the Qubit dsDNA Broad-Range assay kit***
- ***The fragment size is confirmed using the Bio-analyser High Sensitivity DNA Assay***
- ***The KAPA Library Quantification Kit for Illumina® platforms is then used for final quantification prior to denaturation and dilution of the final library.***



Region table for sample 6 : 6. Shantelle 100x

From [bp]	To [bp]	Corr. Area	% of Total	Average Size [bp]	Size distribution in CV [%]	Conc. [pg/µl]	Molarity [pmol/l]	Co lor
200	1,800	498.8	89	493	25.2	581.26	1,853.7	

Steps involved:

Step 7
Sequencing of the final library

1. Denature final pool of DNA library using 0.2N NaOH

2. Dilute quantified final pool of DNA to the desired concentration using Hybridization Buffer (Illumina): Final concentration = 4pM

3. PhiX control is denatured and diluted (to a final concentration of 4pM)
   Control libraries generated from the PhiX virus.
   Characteristics of the PhiX genome provide several benefits:
   Small—PhiX is a small genome, which enables quick alignment and estimation of error rates.
   Diverse—The PhiX genome contains approximately 45% GC and 55% AT.
   Well-Defined—PhiX has a well-defined genome sequence. Illumina cluster generation algorithms are optimized around a balanced representation of A, T, G, and C nucleotides.

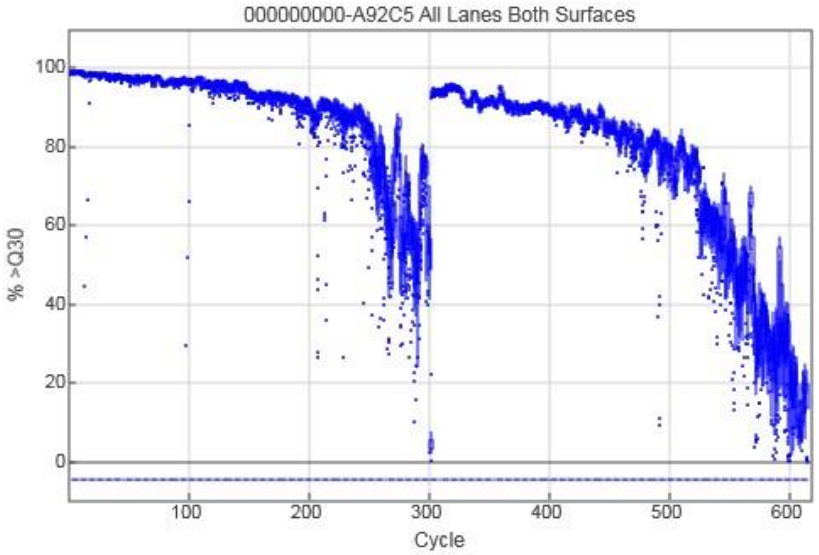
4. Combine sample library and PhiX control:
   For most libraries PhiX is added at 1%. However, for low diversity libraries ≥5% is required.
   We spike PhiX controls at 15%

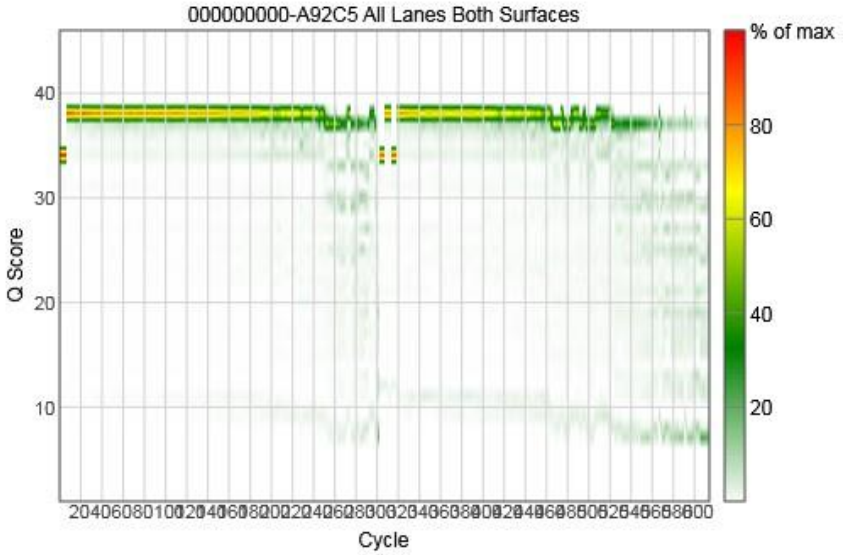
5. Prepare flow cell and kit for the sequencing run

Steps involved:

Step 8
Bio-informatics steps

Quality checks





Bio-informatics pipeline

Summary of 16S rRNA library prep steps involved:

Step 1
Design your experiment:
Controls
DNA extraction protocol
Based on your research question

Step 2
PCR amplification

Step 3
Cleaning of amplicon products

Step 4
Gel electrophoresis and quantification of cleaned amplicon product

Step 5
Agarose gel extraction of pooled product

Step 6
Cleaning of extracted amplicon library and final QC steps

Step 7
Sequencing of the final library

Step 8
Bio-informatics steps