

Genome-resolved metagenomics for tracking antibiotic resistant pathogens in patients

—
Tim Stinear

18th October 2019



Conflicts to declare

None

Imagine this scenario



Jason Kwong



@kwongjc



@kwongjc

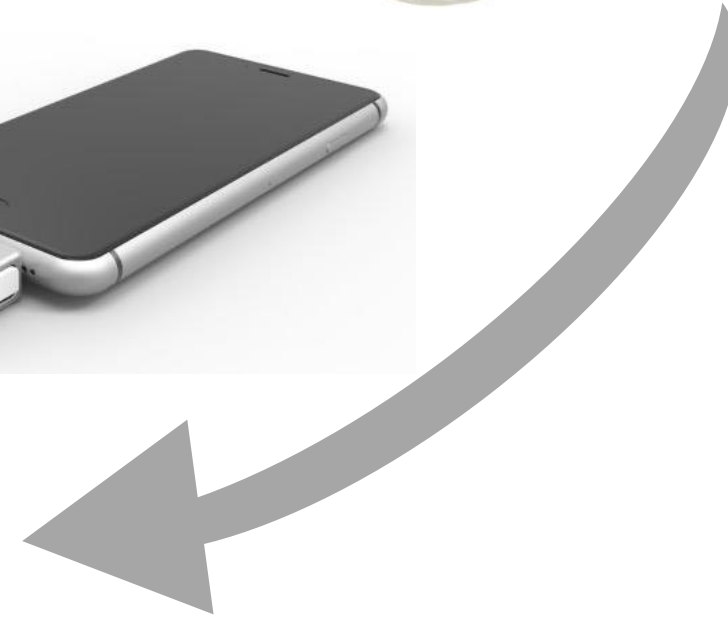


The promise of clinical metagenomics

- Reduce diagnostic delay
- Infer potential antibiotic resistance
- Identify transmission events



Diagnosis



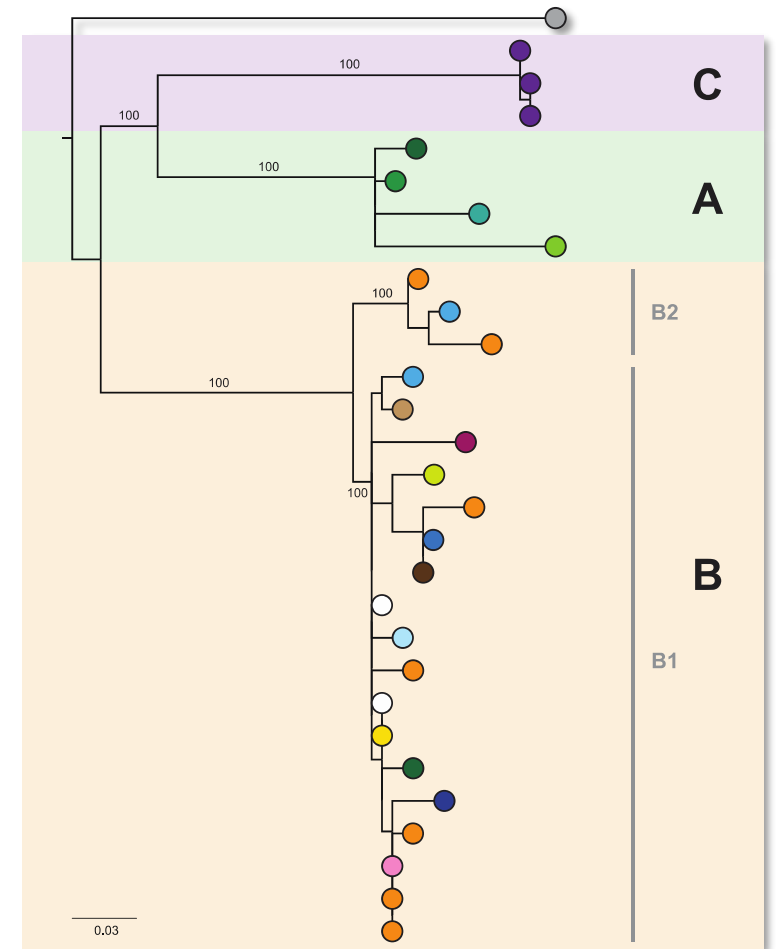
#	Clinical presentation	Sample type	Metagenomics result	Clinical result
1	Eosinophilic meningitis	CSF	No result	Angiostrongylus
2	Meningitis	CSF	S. pneumoniae	S. pneumoniae
3	Meningoencephalitis	CSF	No result	Lupus cerebritis
4	Recurrent brain abscess	CSF	No result	Neuro Behcet's
5	Sepsis	Blood (EDTA)	S. pneumoniae	S. pneumoniae (low level)
6	Sepsis	Blood (EDTA)	N. meningitidis	N. meningitidis
7	Sepsis	Blood (blood culture)	Failed library prep	S. aureus
8	Sepsis	Blood (blood culture)	Failed library prep	E. coli
9	Liver abscess	Aspirate	Entamoeba	Entamoeba
10	Outbreak investigation	Faeces	KPC + VRE	KPC colonisation
11	Outbreak investigation	Rectal swab	KPC	KPC
12	Endocarditis	Aortic valve tissue	No result	Unknown (?M. chimaera)
13	Pancytopenia	Bone marrow aspirate	Leishmania donovani	Leishmania spp.
14	Prosthetic joint infection	Synovial fluid	No result	Ureaplasma
15	Meningoencephalitis	CSF	Pending	Unidentified organism

#	Clinical presentation	Sample type	Metagenomics result	Clinical result
1	Eosinophilic meningitis	CSF	No result	Angiostrongylus
2	Meningitis	CSF	<i>S. pneumoniae</i>	<i>S. pneumoniae</i>
3	Meningoencephalitis	CSF	No result	Lupus cerebritis
4	Recurrent brain abscess	CSF	No result	Neuro Behcet's
5	Sepsis	Blood (EDTA)	<i>S. pneumoniae</i>	<i>S. pneumoniae</i> (low level)
6	Sepsis	Blood (EDTA)	<i>N. meningitidis</i>	<i>N. meningitidis</i>
7	Sepsis	Blood (blood culture)	Failed library prep	<i>S. aureus</i>
8	Sepsis	Blood (blood culture)	Failed library prep	<i>E. coli</i>
9	Liver abscess	Aspirate	<i>Entamoeba</i>	<i>Entamoeba</i>
10	Outbreak investigation	Faeces	KPC + VRE	KPC colonisation
11	Outbreak investigation	Rectal swab	KPC	KPC
12	Endocarditis	Aortic valve tissue	No result	Unknown (? <i>M. chimaera</i>)
13	Pancytopenia	Bone marrow aspirate	<i>Leishmania donovani</i>	<i>Leishmania</i> spp.
14	Prosthetic joint infection	Synovial fluid	No result	Ureaplasma
15	Meningoencephalitis	CSF	Pending	Unidentified organism

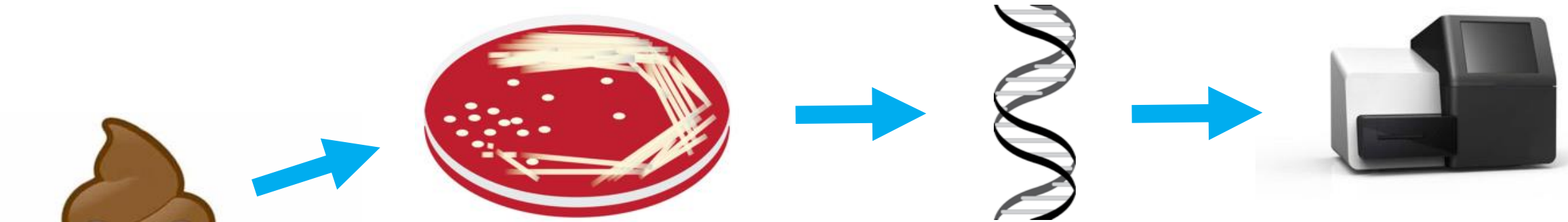
#	Clinical presentation	Sample type	Metagenomics result	Clinical result
1	Eosinophilic meningitis	CSF	No result	Angiostrongylus
2	Meningitis	CSF	S. pneumoniae	S. pneumoniae
3	Meningoencephalitis	CSF	No result	Lupus cerebritis
4	Recurrent brain abscess	CSF	No result	Neuro Behcet's
5	Sepsis	Blood (EDTA)	S. pneumoniae	S. pneumoniae (low level)
6	Sepsis	Blood (EDTA)	N. meningitidis	N. meningitidis
7	Sepsis	Blood (blood culture)	Failed library prep	S. aureus
8	Sepsis	Blood (blood culture)	Failed library prep	E. coli
9	Liver abscess	Aspirate	Entamoeba	Entamoeba
10	Outbreak investigation	Faeces	KPC + VRE	KPC colonisation
11	Outbreak investigation	Rectal swab	KPC	KPC
12	Endocarditis	Aortic valve tissue	No result	Unknown (?M. chimaera)
13	Pancytopenia	Bone marrow aspirate	Leishmania donovani	Leishmania spp.
14	Prosthetic joint infection	Synovial fluid	No result	Ureaplasma
15	Meningoencephalitis	CSF	Pending	Unidentified organism

Genotyping AMR detection

- State-wide outbreak of KPC-producing *Klebsiella pneumoniae*
- Comprised three transmission networks
- Obtained faecal sample from patient implicated in outbreak



Sequencing from culture



Metagenomics



Jason Kwong



Andre Mu

Sequence yield

Illumina

Reads: 22.7M

Yield: 6.4 Gbp

Q score: 33



Detection of antimicrobial resistance genes in Patient A metagenome

Contig	Gene	Gaps	% Coverage	% Identity
FM_1035	blaKPC-2_1	0	100	100

Detection of antimicrobial resistance genes in Patient A metagenome

Contig	Gene	Gaps	% Coverage	% Identity
FM_1035	blaKPC-2_1	0	100	100
FM_452	vanY-B_1	0	100	100
FM_452	vanS-B_1	0	100	99.6
FM_452	vanH-B_1	0	100	99.4
FM_452	vanR-B_1	0	100	99.2
FM_452	vanA-B_1	0	100	98.9
FM_452	vanW-B_1	0	100	97.6
FM_452	vanX-B_1	0	100	96.7

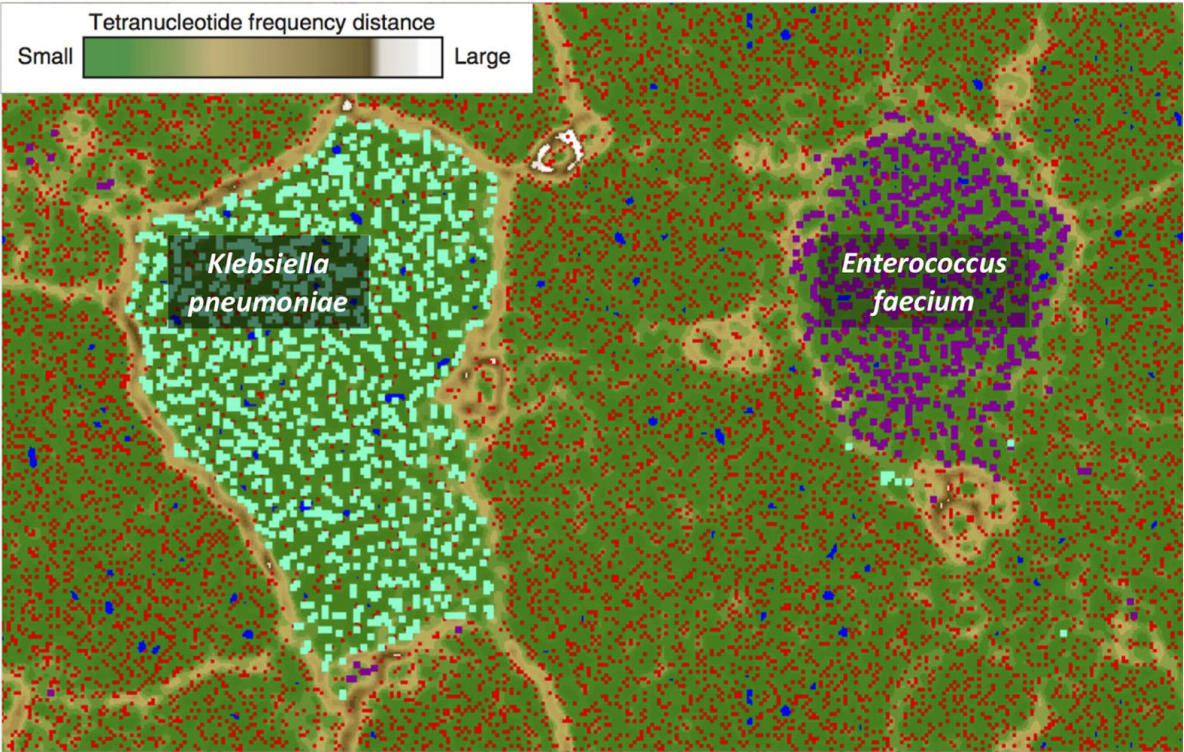
Species ID by 16SrRNA gene reconstruction

Taxon	Length seq DB (bp) ^d	Length inferred seq (bp) ^e	Nucleotide ID ^f	Identity (%)
<i>Akkermansia muciniphila</i>	1,434	1,500	1431	99.8
<i>Bacteroides cellulosilyticus</i>	1,193	1,202	1183	99.2
<i>Bacteroides uniformis</i>	1,145	1,145	1143	99.8
<i>Bacteroides uniformis</i>	680	680	676	99.4
<i>Clostridium bolteae</i>	1,309	1,309	1305	99.7
<i>Clostridium glycyrrhizinilyticum</i>	1,268	1,266	1231	97.1
<i>Dialister invisus</i>	1,290	1,302	1287	99.8
→ <i>Enterococcus faecium</i>	1,344	1,344	1344	100
<i>Eubacterium dolichum</i>	1,484	1,516	1381	93.1
→ <i>Klebsiella pneumoniae</i>	1,161	1,161	1161	100
<i>Lactobacillus pentosus</i>	1,468	1,468	1467	99.9
<i>Parabacteroides merdae</i>	1,376	1,376	1376	100
Uncultured bacterium ^a	1,092	1,092	1092	100
Uncultured organism ^b	1,121	1,121	1121	100
Uncultured organism ^c	1,191	1,191	1191	100

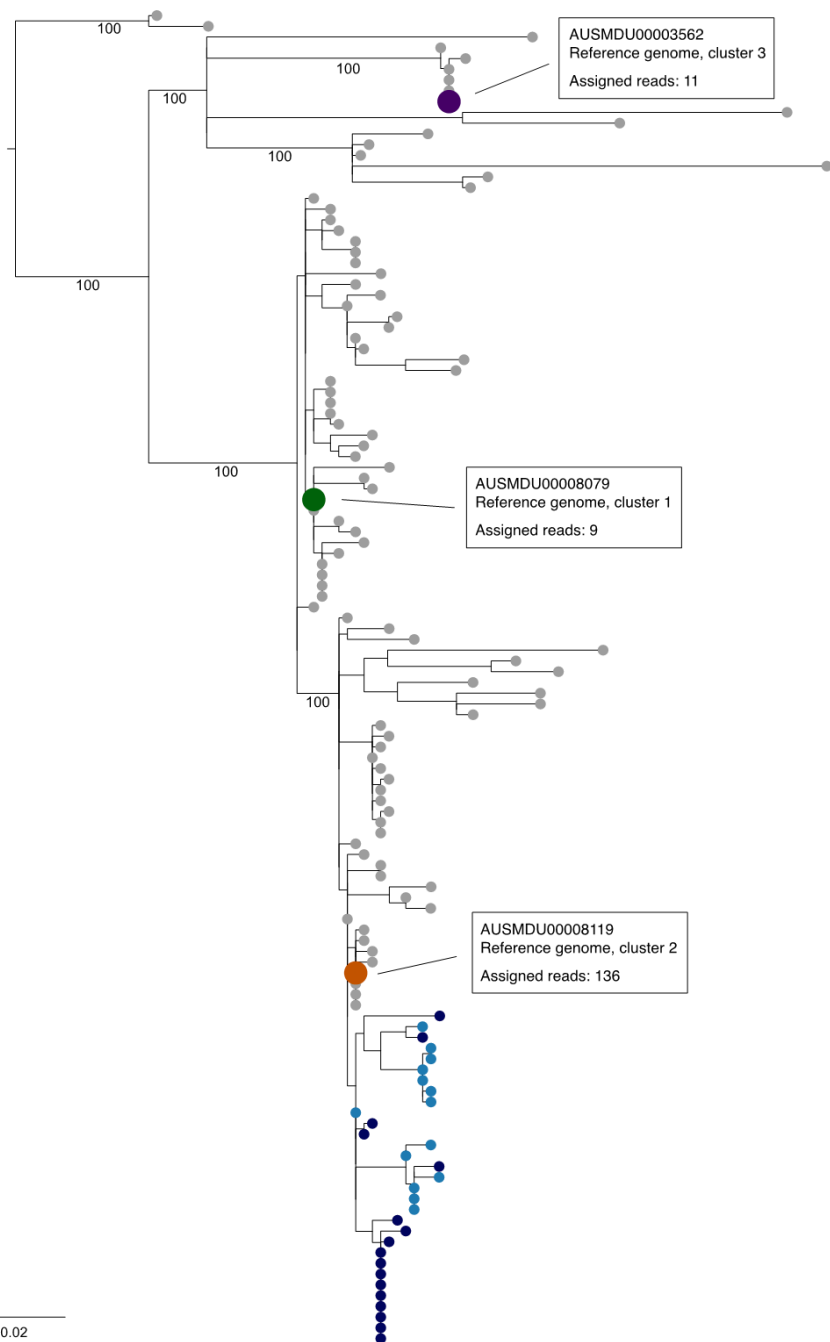
More information?



Binning using tetranucleotide frequency profiles



Sample	Scheme	ST
Patient A faeces	<i>K. pneumoniae</i>	258
Individual colony sequencing	<i>K. pneumoniae</i>	258
Patient A faeces	<i>E. faecium</i>	796
	<i>E. faecium</i>	555
Individual colony sequencing	<i>E. faecium</i>	796

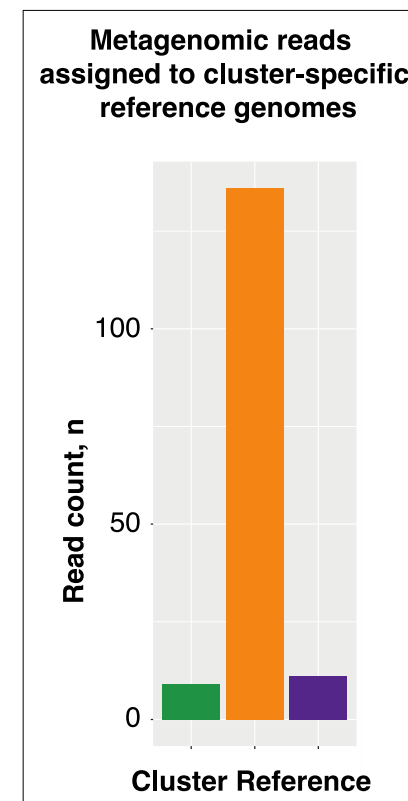


CLUSTER 3

CLUSTER 1

CLUSTER 2

Linked patient to KPC transmission cluster 2



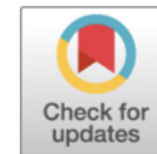


AMERICAN
SOCIETY FOR
MICROBIOLOGY



RESEARCH ARTICLE

Clinical Science and Epidemiology

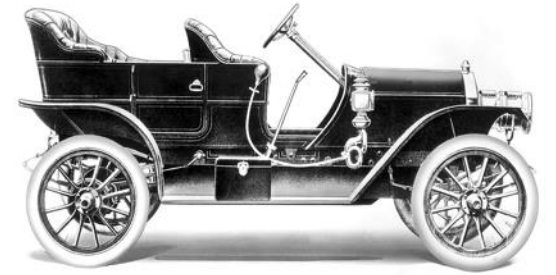


Reconstruction of the Genomes of Drug-Resistant Pathogens for Outbreak Investigation through Metagenomic Sequencing

Andre Mu,^{a,b} Jason C. Kwong,^{a,b,c} Nicole S. Isles,^a Anders Gonçalves da Silva,^{a,b}  Mark B. Schultz,^{a,b} Susan A. Ballard,^{a,b} Courtney R. Lane,^a Glen P. Carter,^b Deborah A. Williamson,^{a,b} Torsten Seemann,^{b,d} Timothy P. Stinear,^b Benjamin P. Howden^{a,b,c}

The Peter Doherty Institute for Infection and Immunity

A joint venture between The University of Melbourne and The Royal Melbourne Hospital



*"The march of progress is neither straight
nor technologically preordained"*

Brandon Keim citing Wells and Norton in:

<http://nautil.us/issue/7/waste/did-cars-save-our-cities-from-horses>



*"The march of progress is neither straight
nor technologically preordained"*

Brandon Keim citing Wells and Norton in:

<http://nautil.us/issue/7/waste/did-cars-save-our-cities-from-horses>

Assessing the impact of long reads



Assessing the impact of long reads



- DNA extraction (24h)
 - 3 M guanidine HCl, 20% Tween-20
 - Qiagen Genomic-tips
 - Purification with Ampure XP beads
 - Size selection on Bluepippin
- Library prep/sequencing (48h)
 - Ligation Sequencing Kit 1D (SQK-LSK109)
 - Prime and load R9.4 flow cell (FLO-MIN106)
 - Sequencing on MinION



Koen Vandelannoote



Assessing the impact of long reads

- Data analysis (>48h):
 - Guppy-GPU v3.2.4: basecalling on HPC-GPU node (12h)
 - Porechop v0.2.4: Adapters trim, removing hybrid reads (1h)
 - Filtlong v0.2.0: filtering ONT reads
 - Minimap2 v2.17: Removing human reads (30min)
 - Flye v2.6 : De novo assembly (36h)
 - Racon v1.3.2: long read polishing (2 rounds) (1h)
 - Medaka v0.10.0: long read polishing contigs using basecaller
“error-profile aware” model (30 min)



Koen Vandelannoote



Assessing the impact of long reads

Conclusion Part I

At least 5 days for first results from receipt of fecal specimen.

Assessing the impact of long reads

Sequence yield



Illumina:

Reads: 22.7M

Yield: 6.4 Gbp

Q score: 33

ONT:

Reads: 4.7M

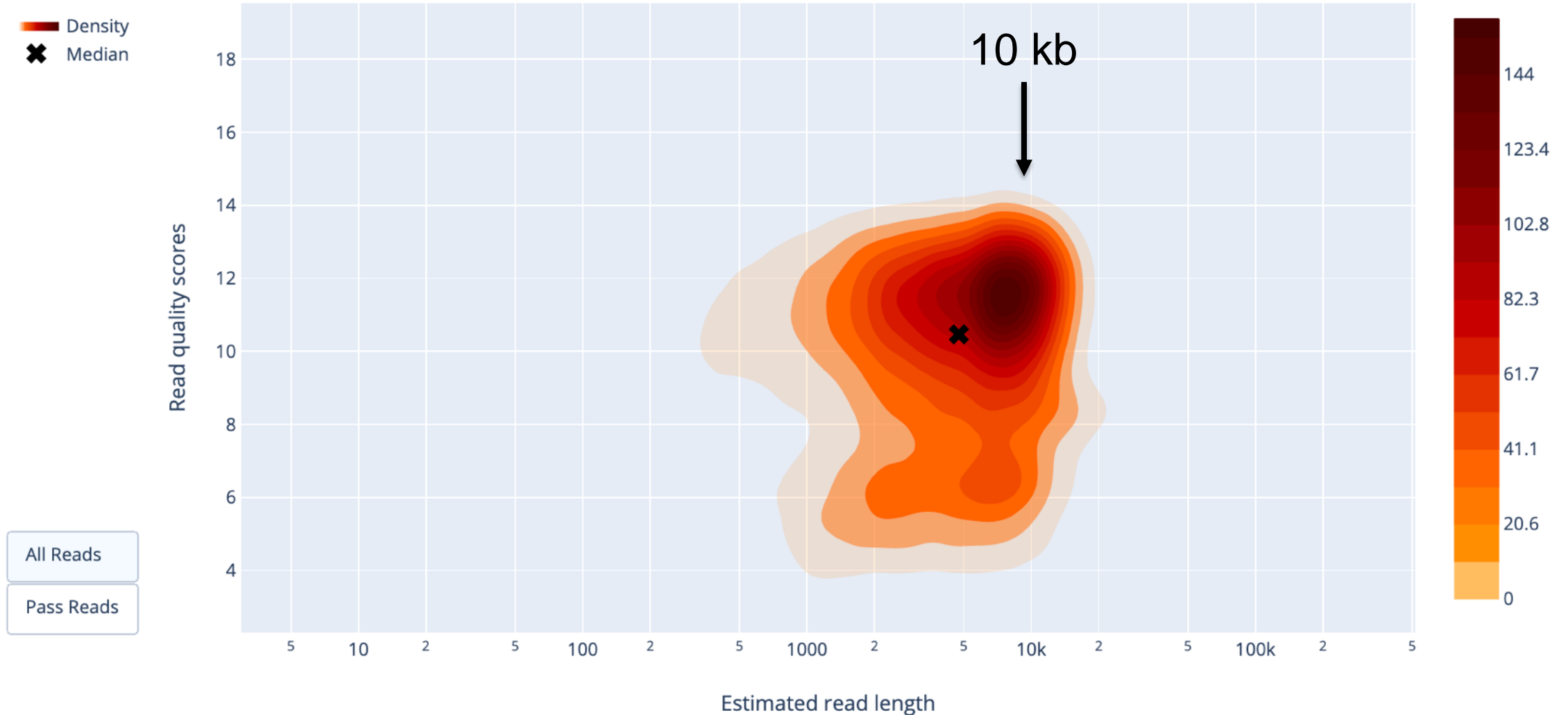
Yield: 31.8 Gbp

Q score: 18



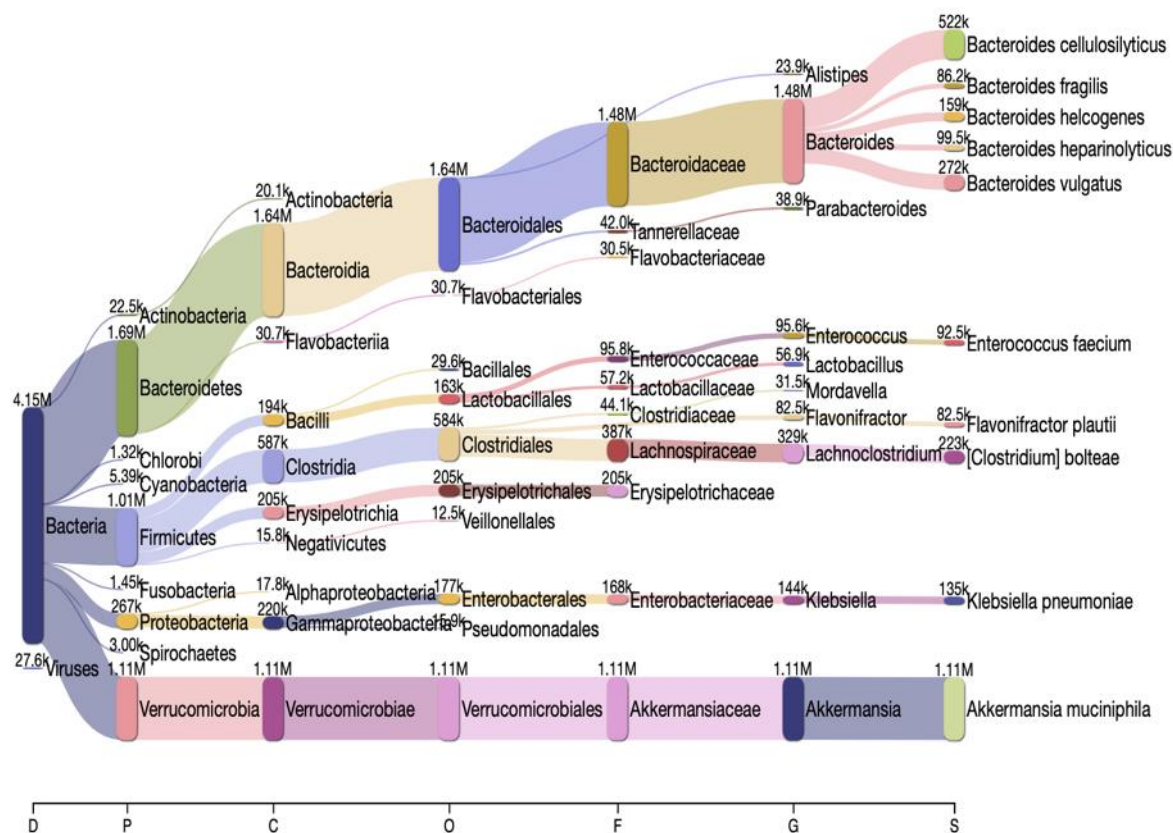
Assessing the impact of long reads

Nanopore: Lower quality reads, but loooooong



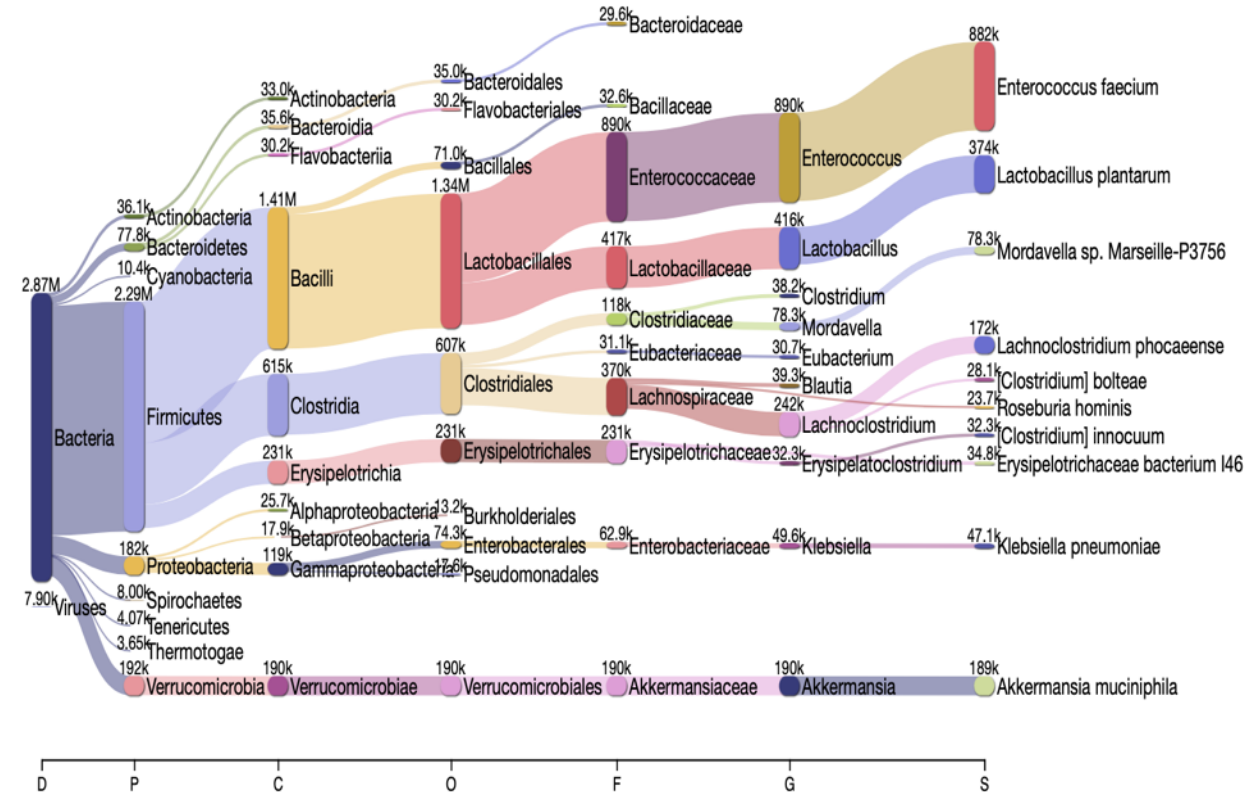
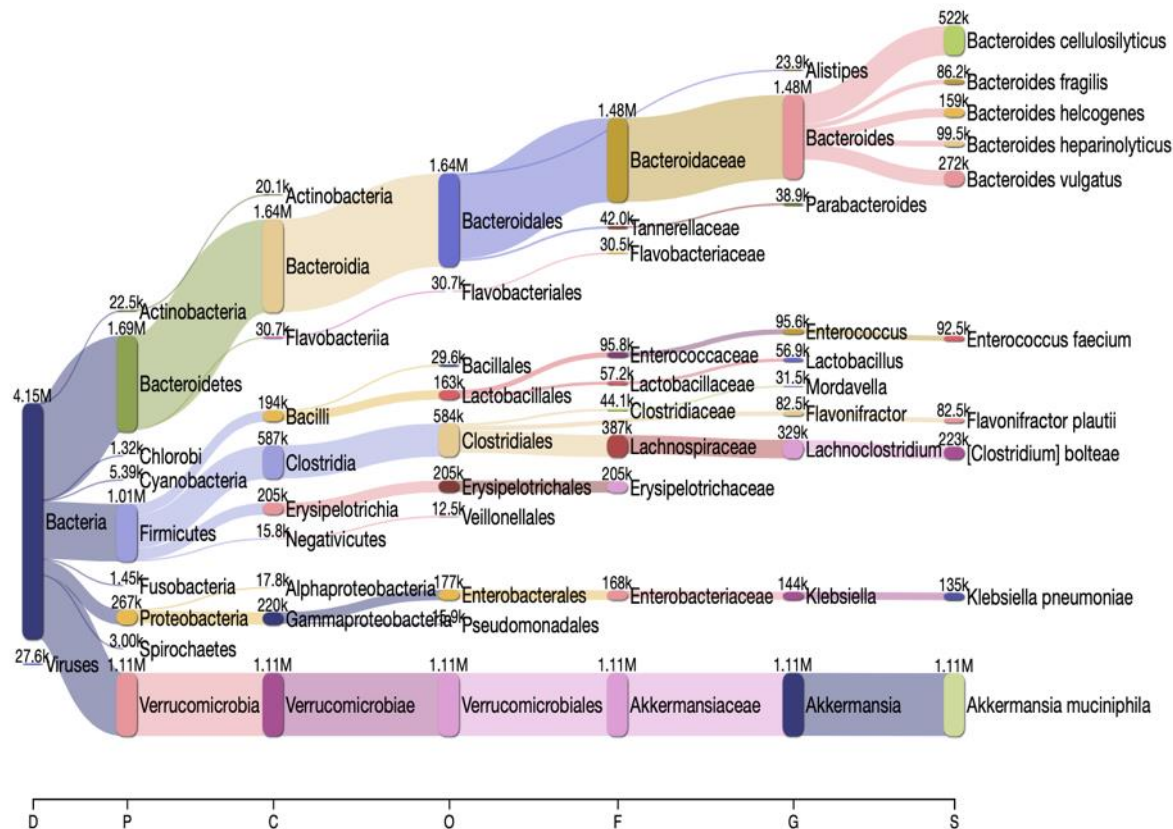
Breadth of coverage similar, proportions bacteria different

illumina®



Breadth of coverage similar, proportions bacteria different

illumina®



Nanopore provides greater genome coverage



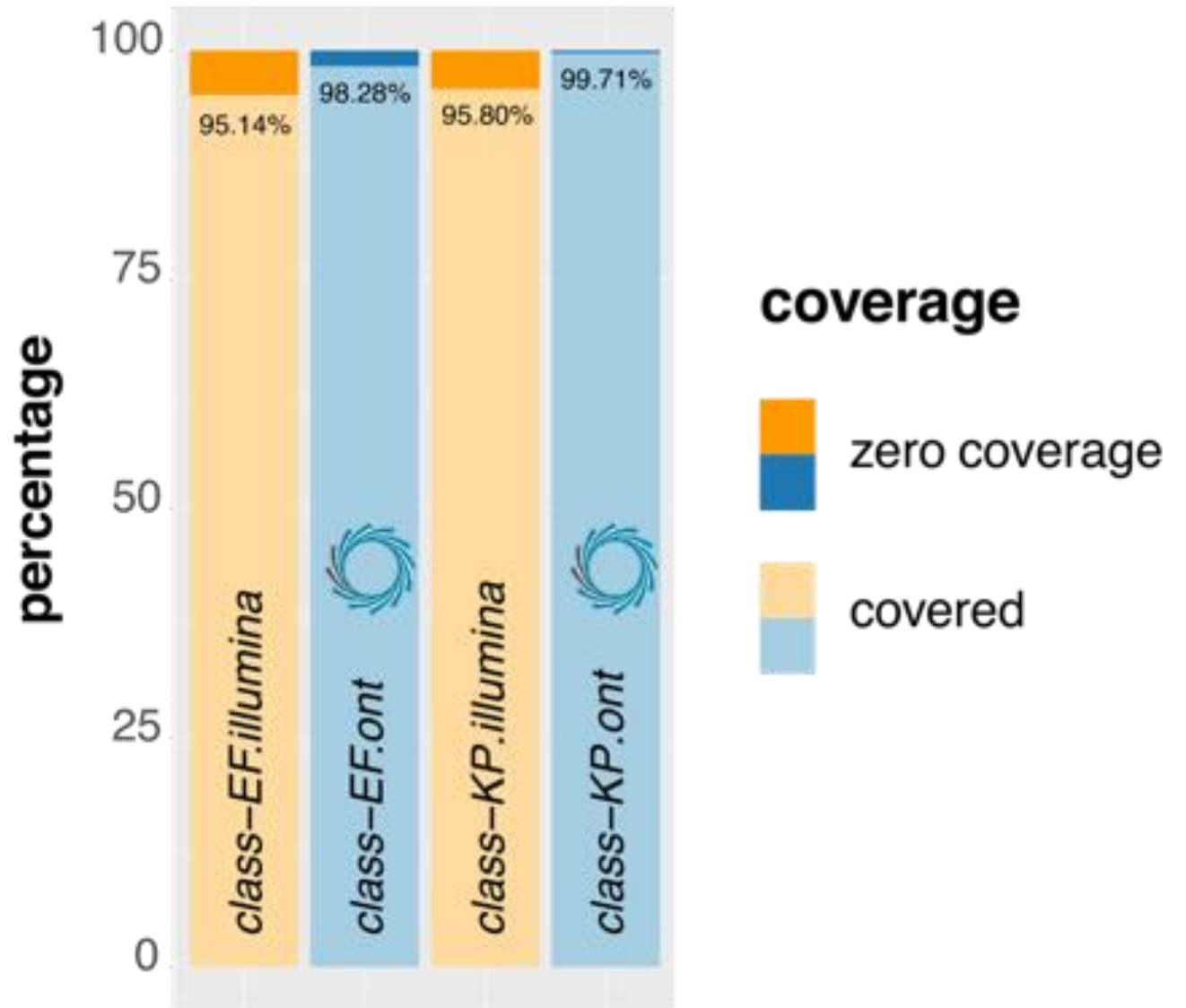
De novo assembly comparisons

Sample	No. contigs	Base pairs	Min length	Avg length	Max length	N50
Illumina	113,623	106,938,398	200	941	203,719	4,750
Nanopore	1,758	87,265,890	257	49,639	5,401,097	96,648



Nanopore provides greater genome coverage

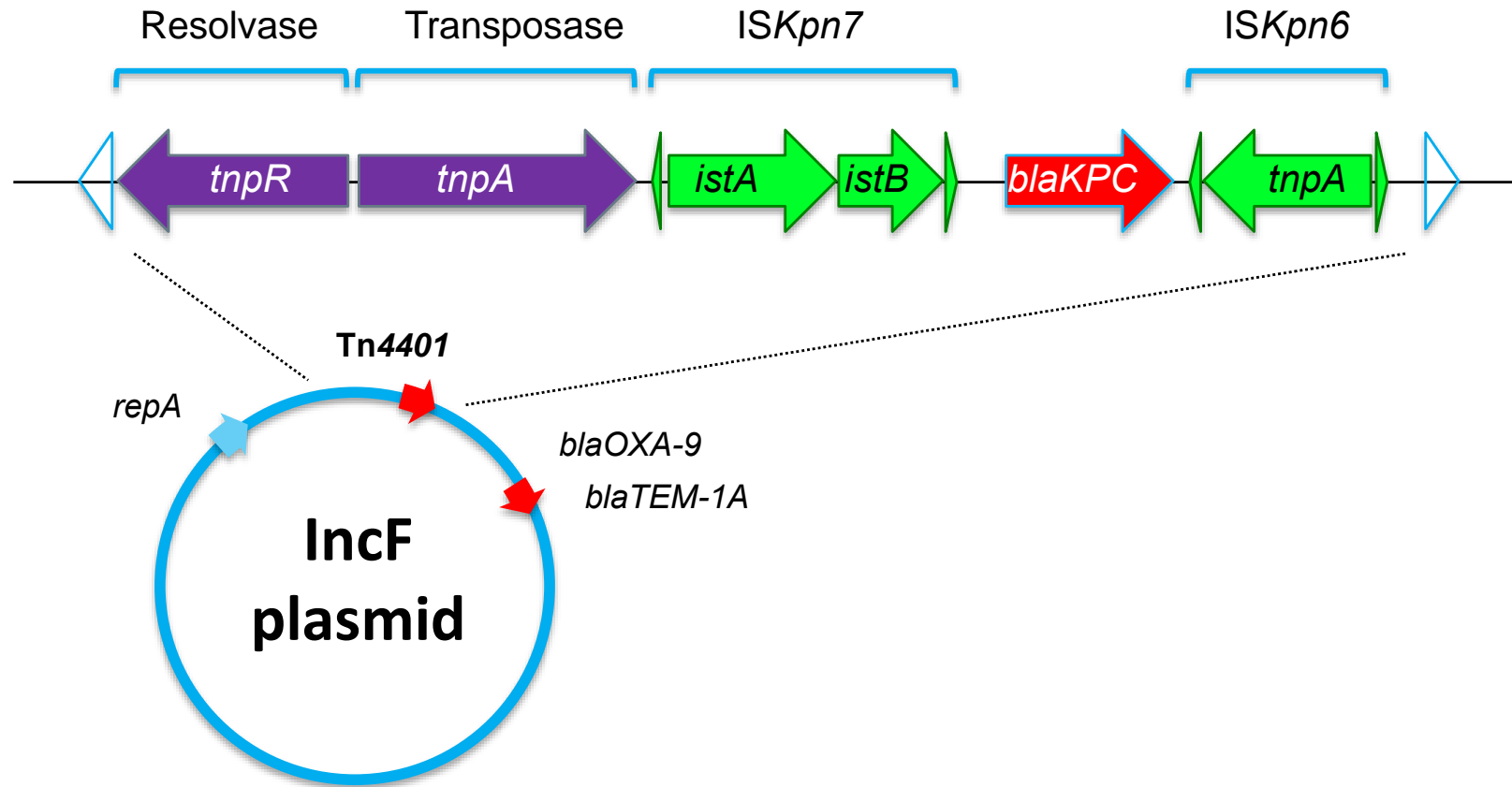
Proportion of reference genome covered by metagenomic reads



Nanopore provides greater genome coverage



Tn4401 – IncF plasmid



Gene	Resistance	Illumina Coverage %	Illumina Identity %	Nanopore Coverage %	Nanopore Identity %
blaKPC-2	CARBAPENEM	100.00	100.00	100.00	99.77
blaSHV-158	BETA-LACTAM	-	-	100.00	99.88
blaTEM-1	BETA-LACTAM	-	-	20.09	100.00
blaOXA-9	BETA-LACTAM	100.00	99.88	98.67	98.55
aac(6')-Ib	AMIKACIN/KANAMYCIN	-	-	100.00	99.28
aph(3')-Ia	KANAMYCIN	100.00	100.00	99.02	98.30
aadA2	STREPTOMYCIN	83.33	99.39	94.32	94.37
dfrA12	TRIMETHOPRIM	100.00	99.80	97.59	96.79
sul1	SULFONAMIDE	100.00	100.00	99.05	98.33
mph(A)	MACROLIDE	18.78	99.42	99.24	98.27
oqxA	PHENICOL/QUINOLONE	100.00	100.00	100.00	99.58
oqxB	PHENICOL/QUINOLONE	100.00	100.00	99.97	99.59
fosA6	FOSFOMYCIN	100.00	98.81	100.00	98.81
catA1	CHLORAMPHENICOL	-	-	98.64	98.33

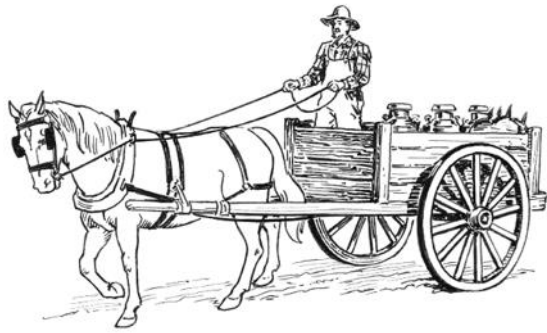
Gene	Resistance	Illumina Coverage %	Illumina Identity %	Nanopore Coverage %	Nanopore Identity %
blaKPC-2	CARBAPENEM	100.00	100.00	100.00	99.77
bla					
bla					
bla					
bla					
aa					
ap					
aa					
dfr					
sul1	SULFONAMIDE	100.00	100.00	99.05	98.33
mph(A)	MACROLIDE	18.78	99.42	99.24	98.27
oqx A	PHENICOL/QUINOLONE	100.00	100.00	100.00	99.58
oqx B	PHENICOL/QUINOLONE	100.00	100.00	99.97	99.59
fosA6	FOSFOMYCIN	100.00	98.81	100.00	98.81
catA1	CHLORAMPHENICOL	-	-	98.64	98.33

- Nanopore 13/14 genes >90% coverage vs Illumina 8/14
- Illumina more often higher percentage nucleotide identity

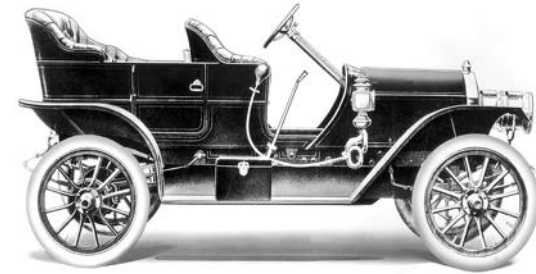
Conclusions Part II

- Need longer, higher quality reads
 - Define sequencing effort required
- Need better DNA extraction methods
- Sequencing costs too high
- Data analysis times too long

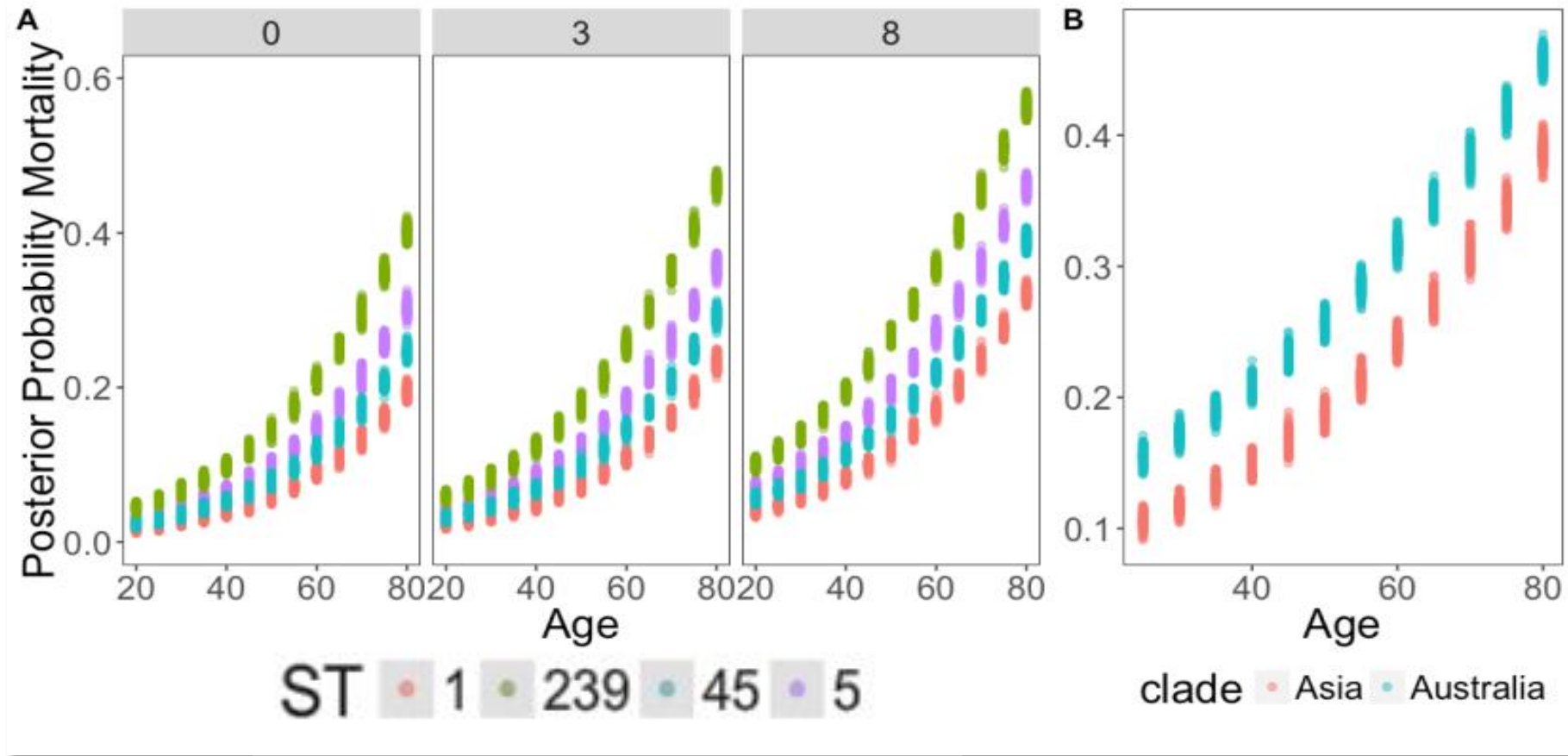




?



S. aureus genotype predicts mortality risk



Gonçalves da Silva et al, unpublished

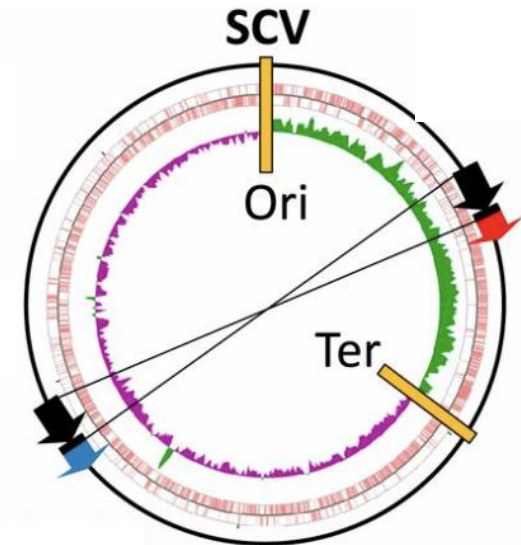
The Peter Doherty Institute for Infection and Immunity
A joint venture between The University of Melbourne and The Royal Melbourne Hospital

S. aureus chromosome rearrangement, invasive disease

PNAS

Unstable chromosome rearrangements in *Staphylococcus aureus* cause phenotype switching associated with persistent infections

Romain Guérillot^a, Xenia Kostoulas^b, Liam Donovan^a, Lucy Li^a, Glen P. Carter^{a,c}, Abderrahman Hachani^a, Koen Vandelannoote^a, Stefano Giulieri^a, Ian R. Monk^a, Mayu Kunimoto^d, Lora Starrs^d, Gaétan Burgio^d, Torsten Seemann^{c,e}, Anton Y. Peleg^{b,f}, Timothy P. Stinear^{a,c,1}, and Benjamin P. Howden^{a,c,e,g,1,2}



How far to go before metaG is a routine test?

- Rapid, cost-effective diagnosis
- Risk profiling (max diagnostic yield)
 - ARD presence/context
 - pathogen genotype
 - specific chromosome conformations
- Informed treatment



Diagnosis



Acknowledgements



- **MDU PHL: Benjamin Howden**, Deborah Williamson, Susan Ballard
- **Microbiology and Immunology: Jason Kwong, Andre Mu, Koen Vandelannoote, Nicole Isles, Mark Schultz, Sara Zufan**, Glen Carter, Andrew Buultjens, Sarah Baines, George Tairoa
- **MDU PHL bioinformatics team: Torsten Seemann; Anders Gonçalves Da Silva**;
- **MDU PHL Enteric Laboratory** (Mary Valcanis; Jess Barnden; Nicole Orlando)
- **MDU PHL Molecular Unit** (Michelle Sait; Takehiro Tomita; Karolina Mercoulia)
- **MDU PHL Epidemiology** (Janet Strachan; Courtney Lane)
- **MDU PHL AMR Unit** (Norelle Sherry, Kerrie Stevens)



Back to our imagined scenario



