

Sampling and sequencing

Dr Josh Quick
Sr Research Fellow, University of Birmingham

Introduction to today

9.30 Introduction to sequencing, extraction and library preparation

10.30 Coffee break

11.00 Introduction to long-read metagenomics

12.00 Lunch

13:30 Library preparation, flowcell QC and loading

14:30 Coffee break

15.00 Running MinKNOW, basecalling, bioinformatics

2001

Microbe >\$1,000,000 and 3 years

First human genome \$2,700,000,000 and 13 years
(\$100,000,000 in reagents)



2012

Microbial genome: \$50 and 2 days

Human genome: \$1000-\$10000 in a week



2015

Portable, real-time sequencing



Nanopore development



Nick Loman

@pathogenomenick

Following

A genuine >8kb Oxford @nanopore read off our instrument THIS MORNING. And happily it's actually useful for diagnosis!



A *P. aeruginosa* serotype-defining single read from our first ...

Here is, I think, the first publically-available Oxford Nanopore read to be published. This came off our MinION instrument this morning (Wednesday 11th June). The DNA was derived from Ps...
figshare.com

11:40 PM - 10 Jun 2014

119 Retweets 43 Likes



10

119

43



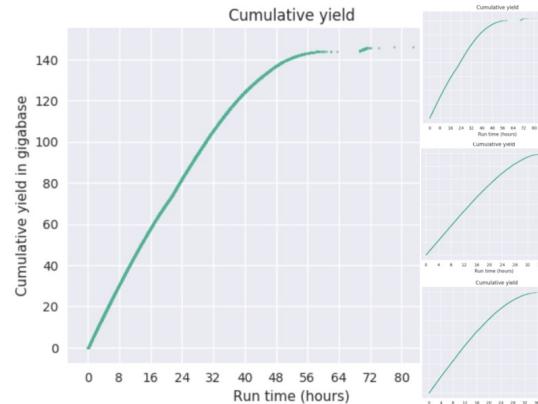
Nick Loman
@pathogenomenick

Following

New Zymo long-read nanopore mock community data available using new MinKNOW ACS mode: matched PromethION (146, 148Gb) and GridION (14, 16Gb) runs for log and even community samples now with batch numbers!

github.com/LomanLab/mockc...

Thanks to @Scalene @samstudio8 !
Preprint soon.

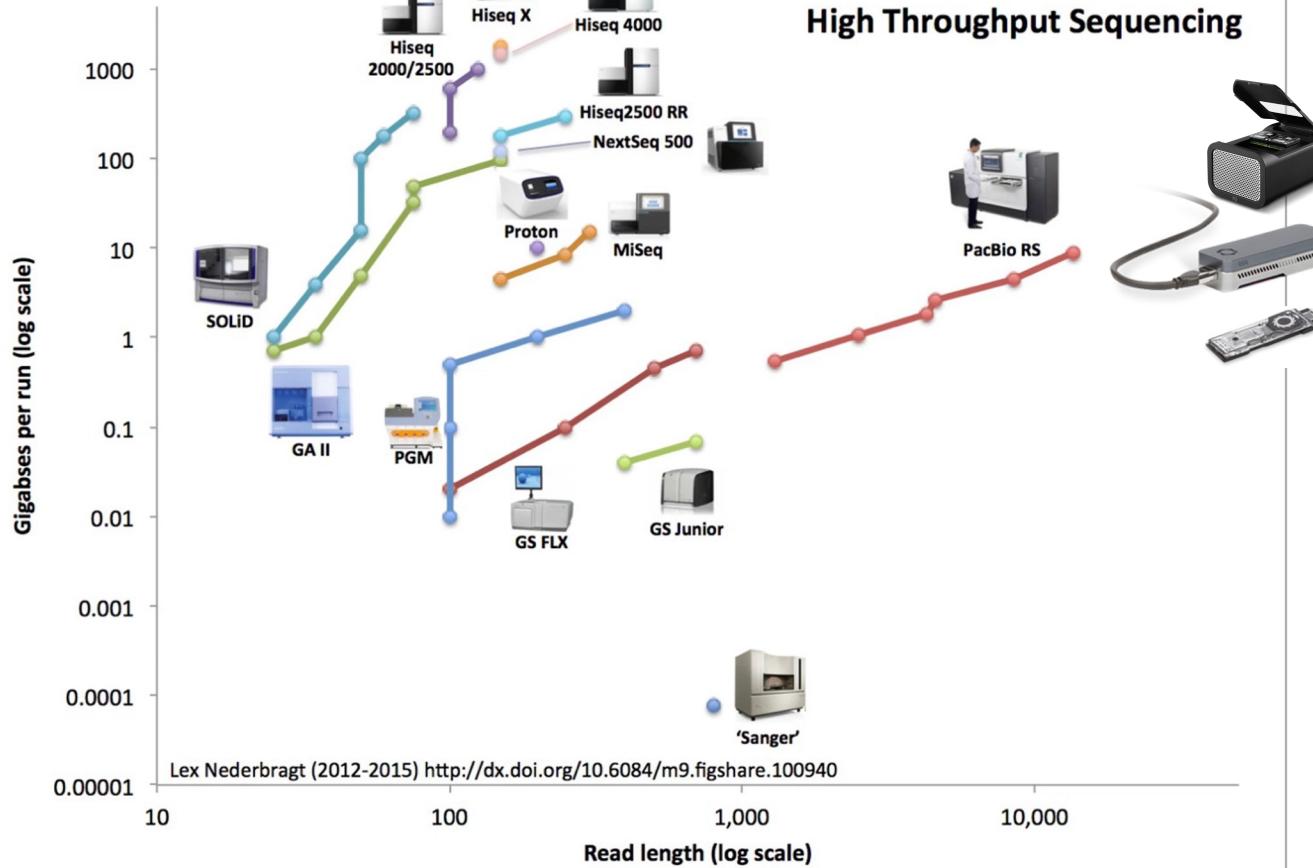


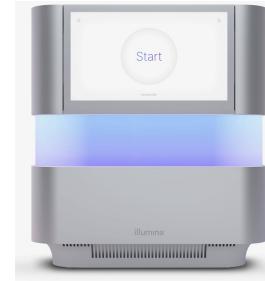
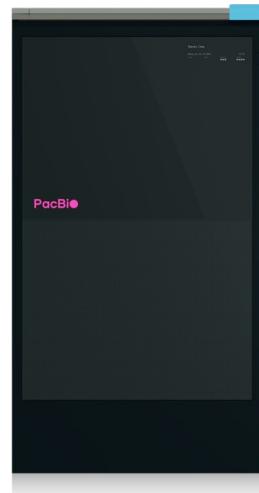
4:26 AM - 17 Oct 2018

32 Retweets 70 Likes



Developments in High Throughput Sequencing





Factory style

Benchtop/mid-range

Portable



~3.6Tb/64h



6Tb/44h



15Gb/4h

Factory style

100Gb/48h



1-2Gb/24h



360Gb/48h



20Gb/48h



50Gb/12h

Benchtop/mid-range



200Gb/48h

Portable



\$150,000



\$1,000,000

Factory style



\$800,000

\$50,000



\$20,000



\$300,000



\$1000



\$50,000



\$10,000



Benchtop/mid-range

Portable

Nanopore flowcells



	Channels	Yield (Gb)	Cost
Flongle	126	<1	£736 (12)
MinION	2,000	10-20	£430 (96) - £810 (1)
PromethION	12,000	100-200	£540 (2880) - £810 (4)

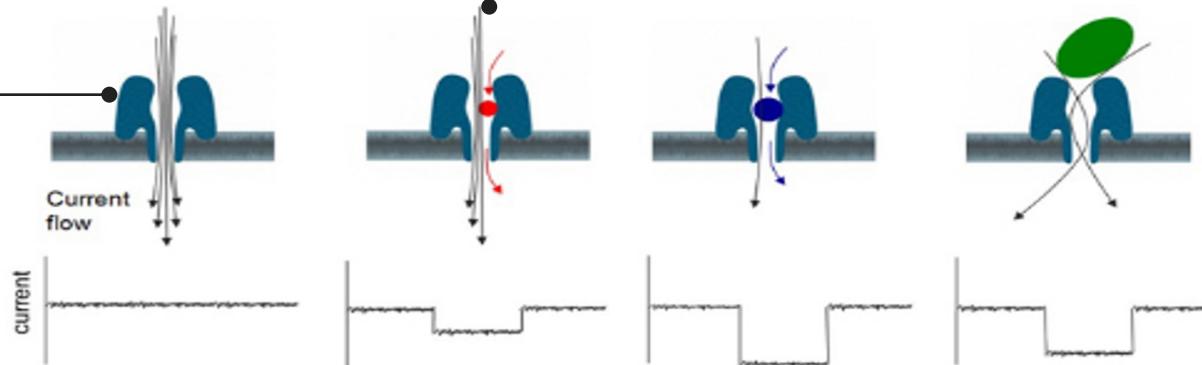
Nanopore detection

1

Nanopore creates hole in membrane
Current passes through nanopore

2

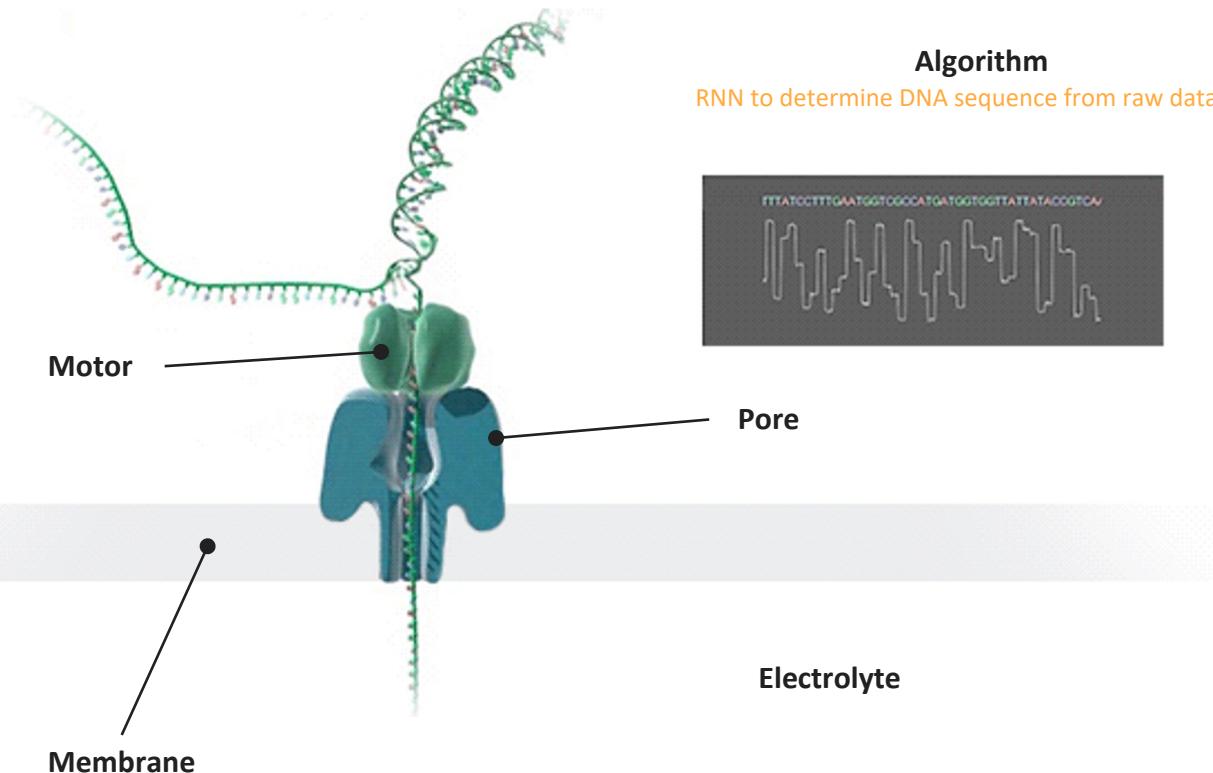
As analyte passes through or near the nanopore,
this creates characteristic disruptions in the current



3

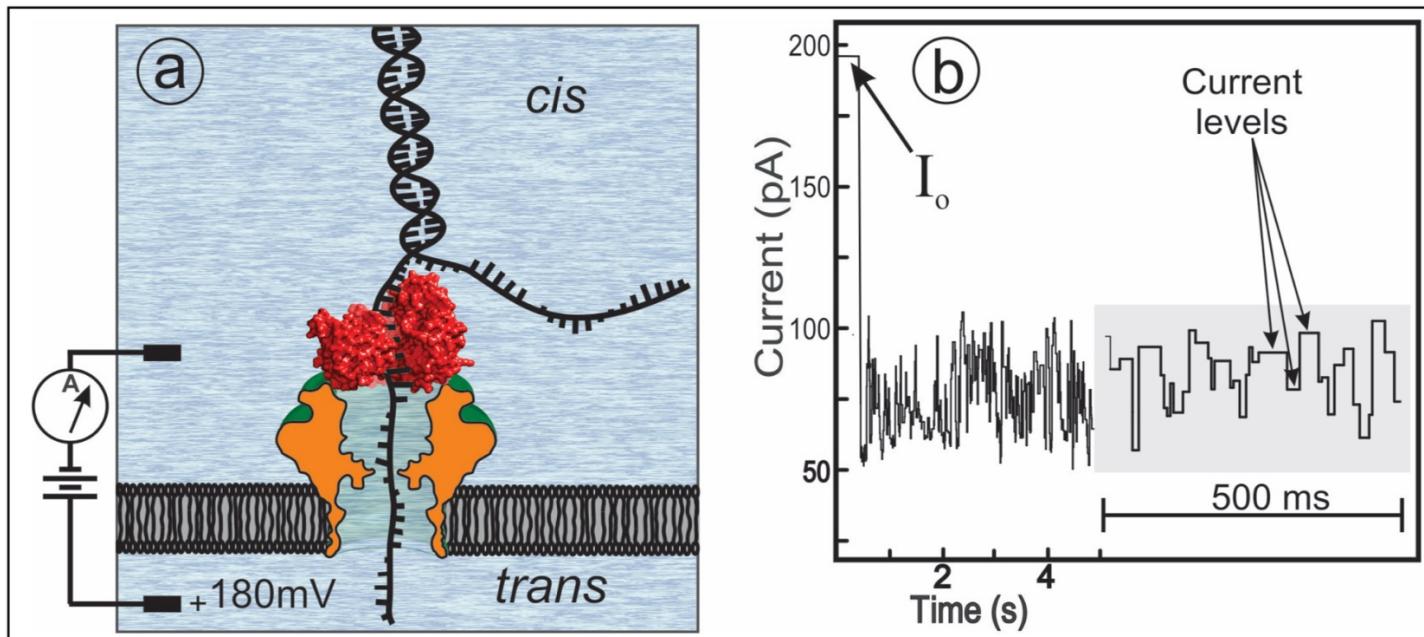
Current disruption is interpreted to understand the
identity of the analyte

Strand sequencing



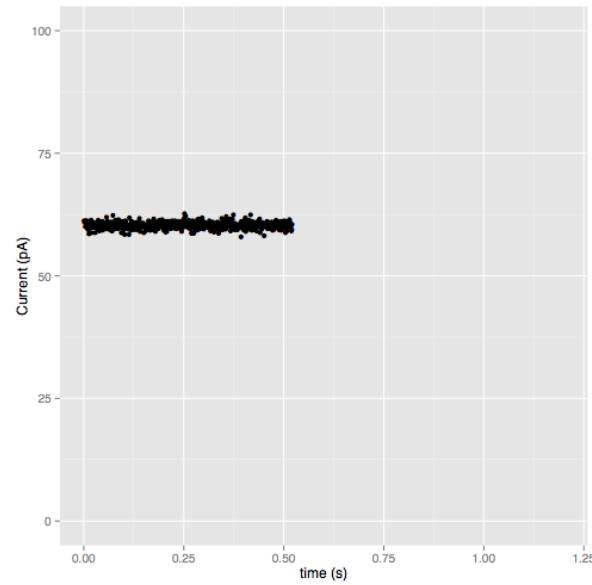


Nanopore Sequencing

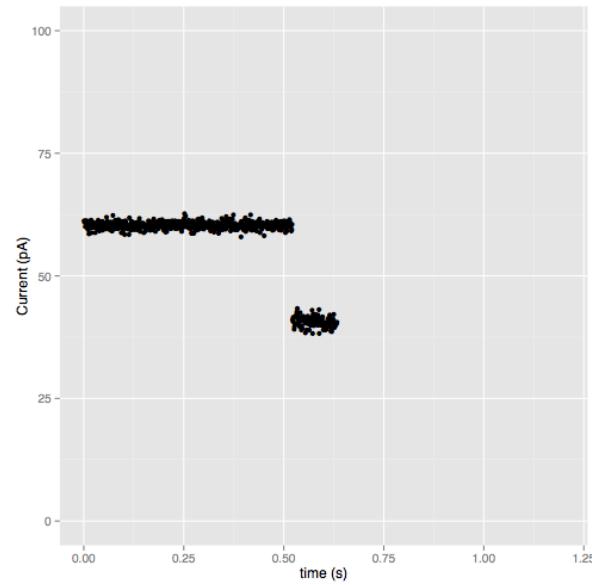


Understanding Nanopore Sequencing: A guide to DNA and RNA analysis, Deamer and Branton

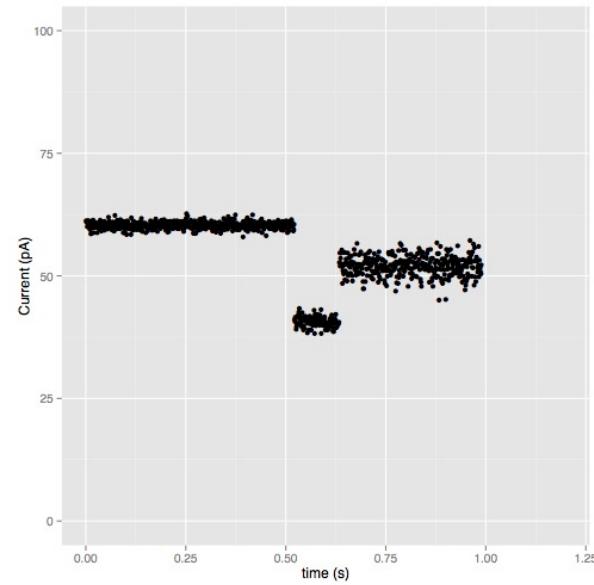
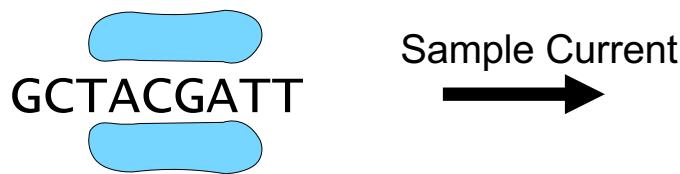
Nanopore Sequencing



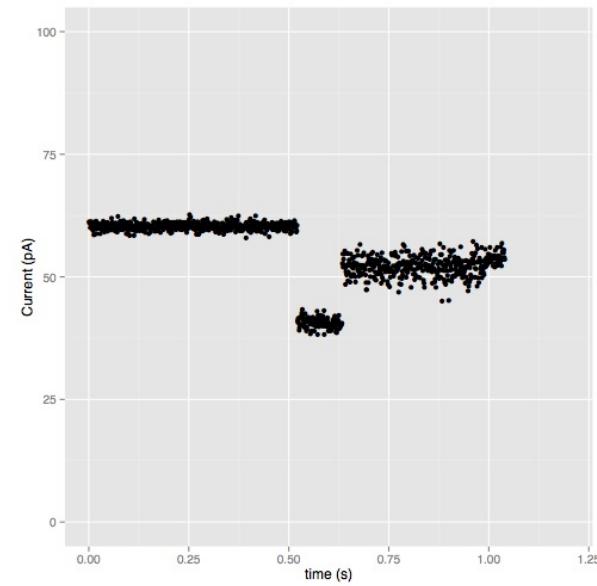
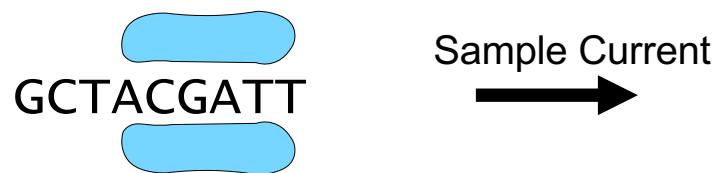
Nanopore Sequencing



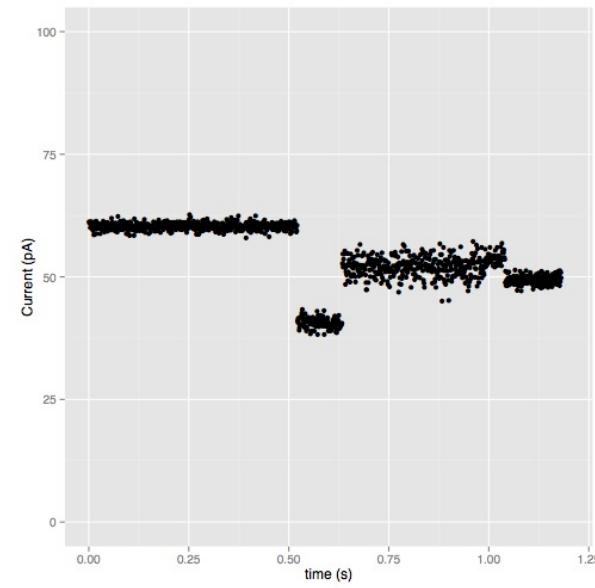
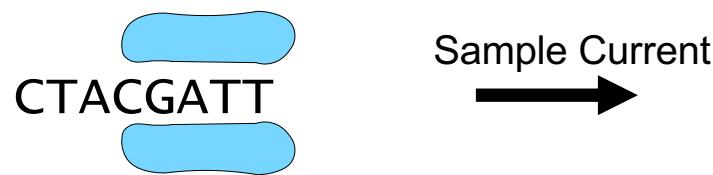
Nanopore Sequencing

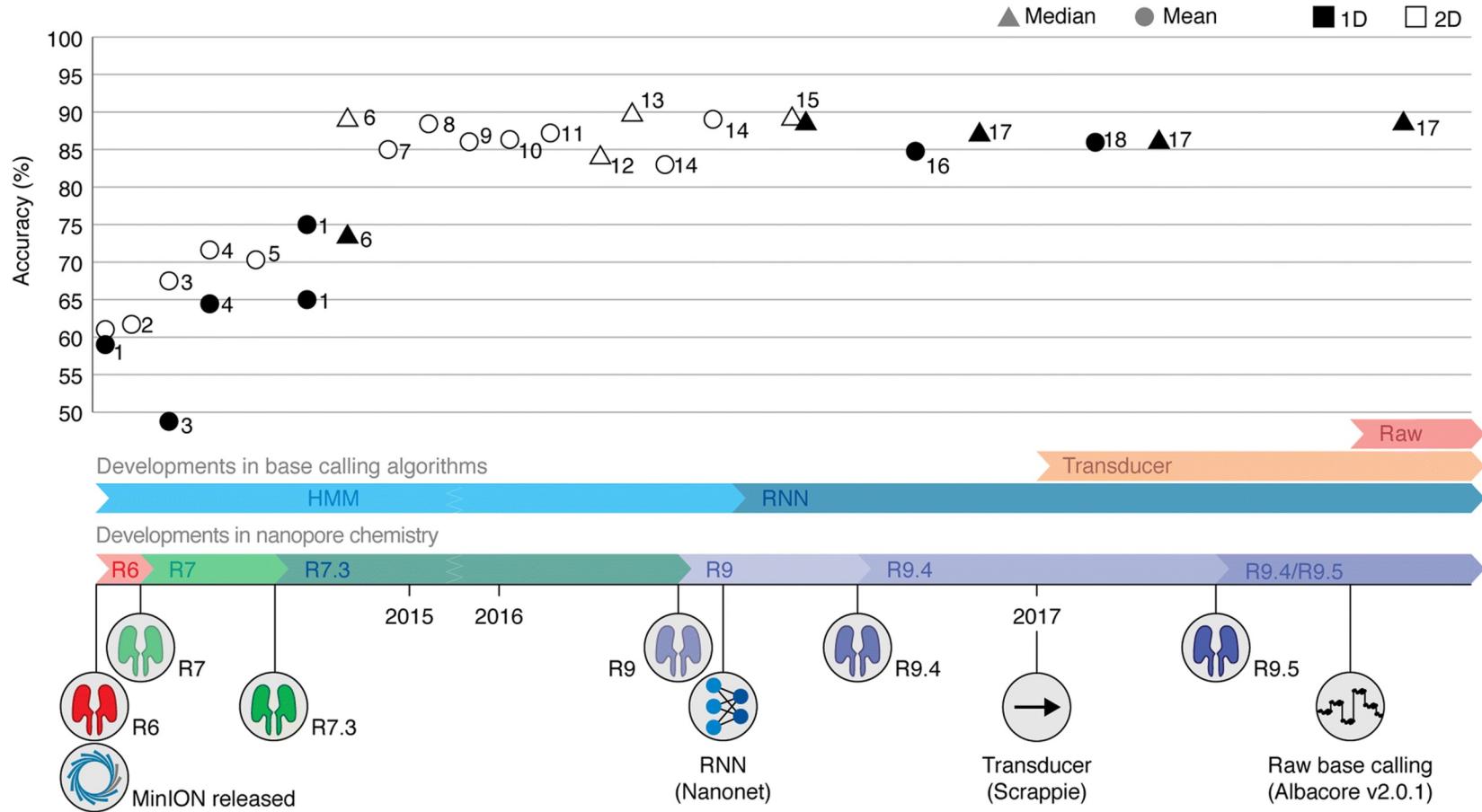


Nanopore Sequencing



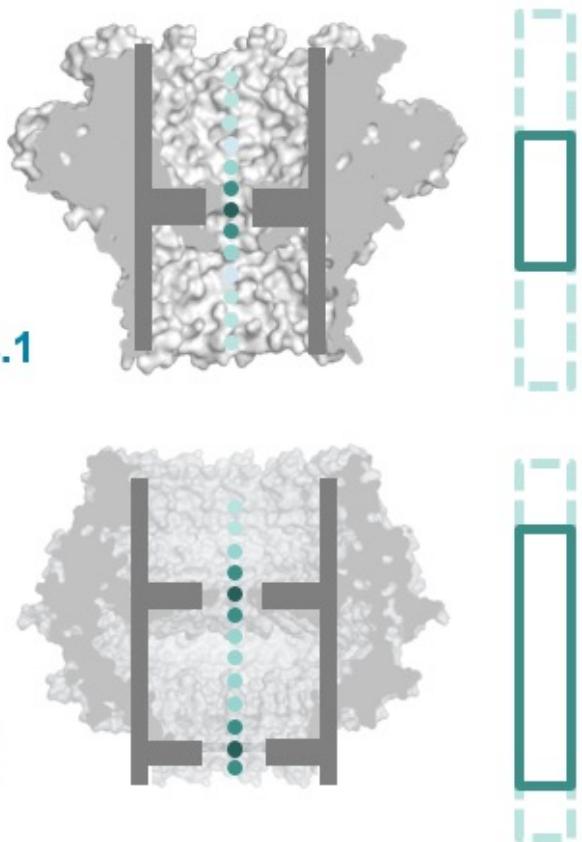
Nanopore Sequencing





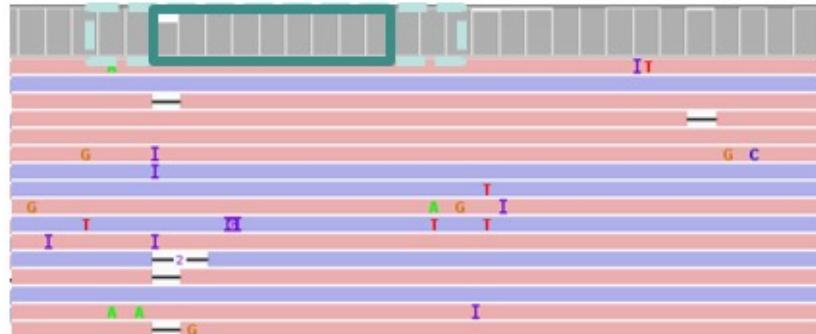
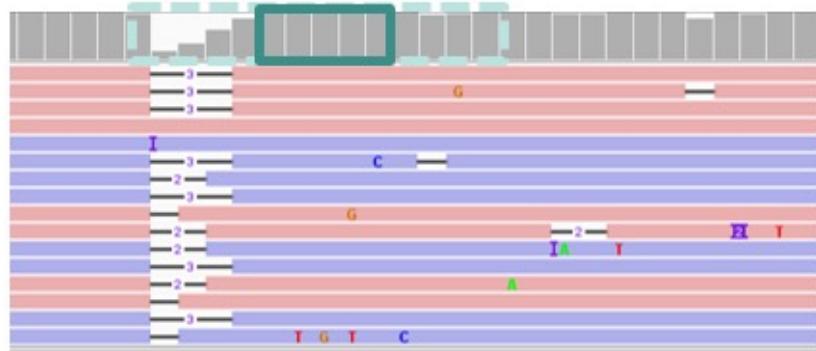
Current chemistry

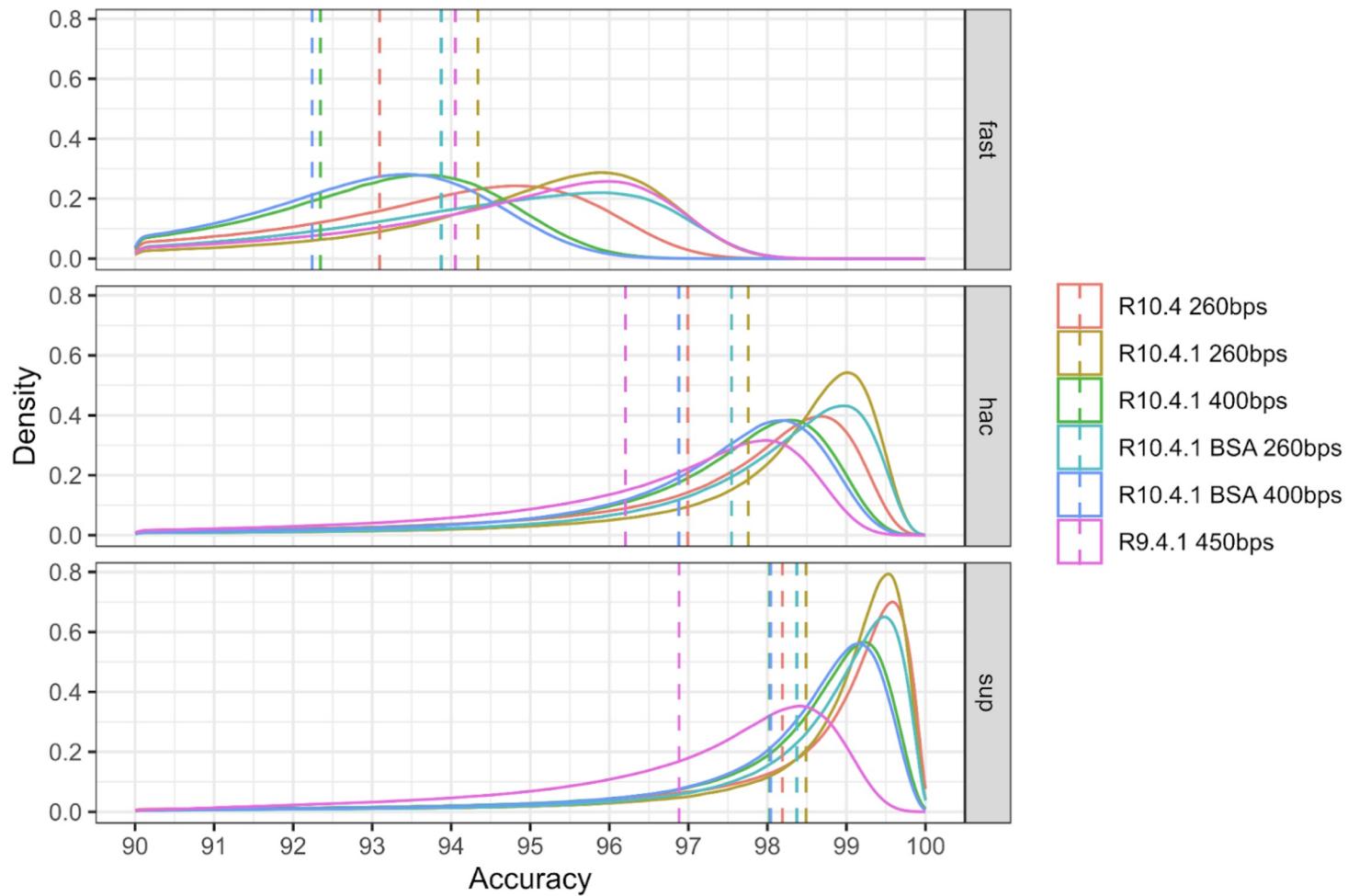
R9.4.1

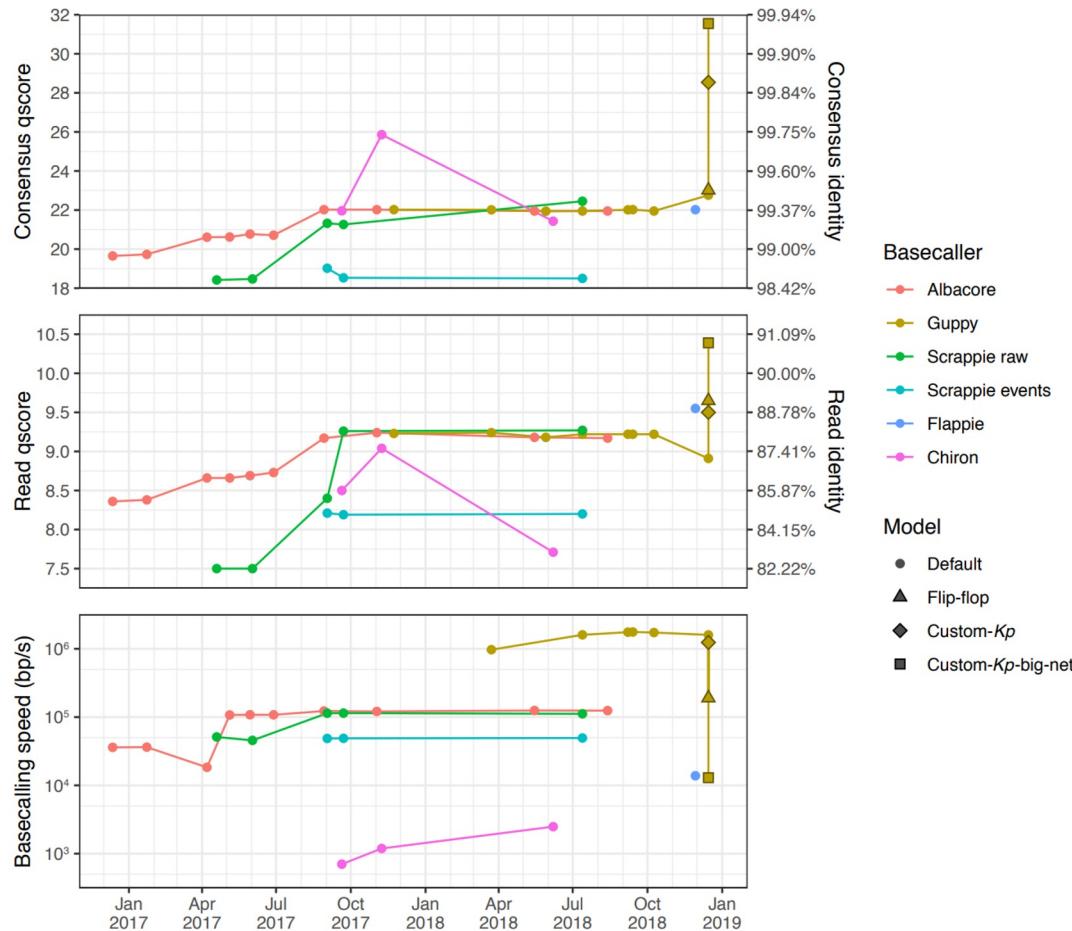


R10

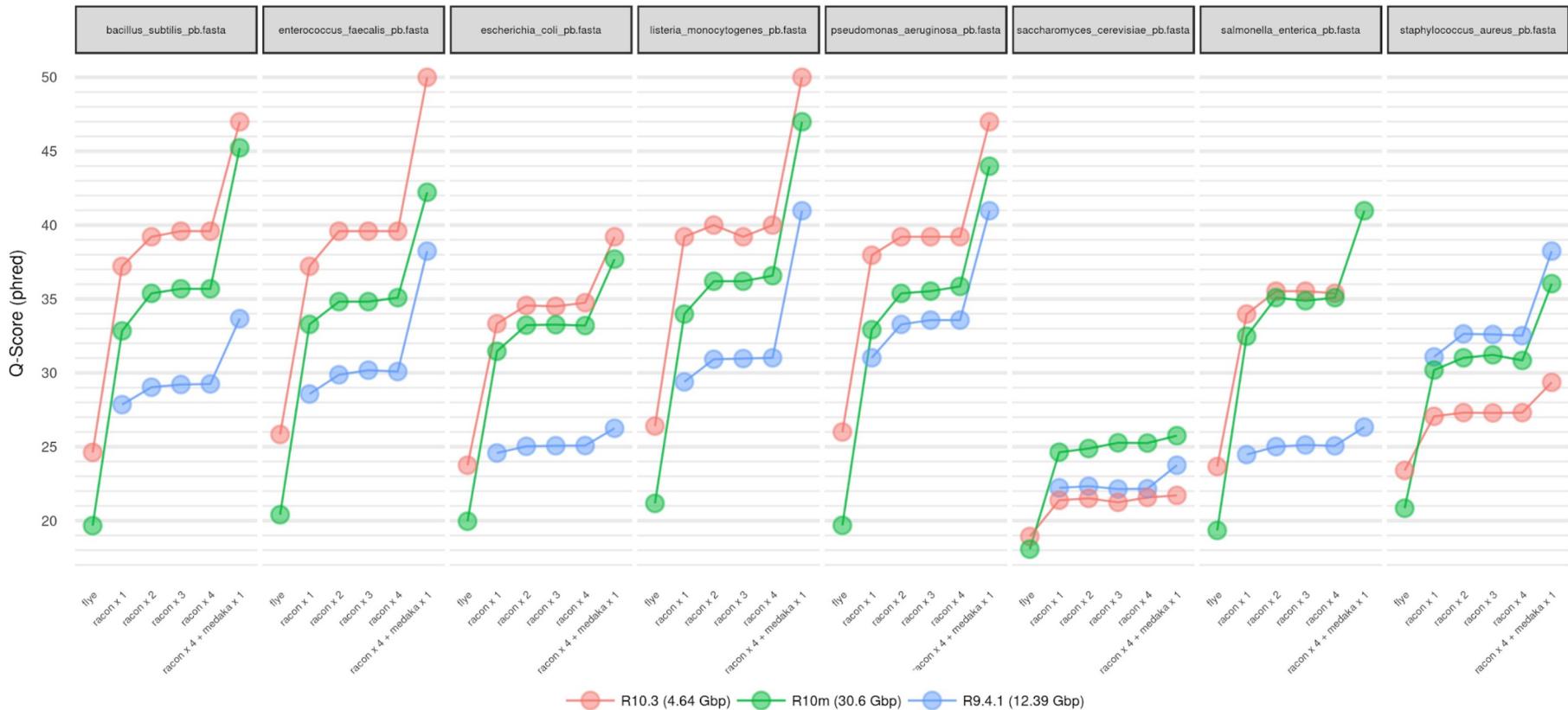
ATCGGGAAAAAAATCACGCCACGTCCAAA







Polishing



Note: the difference in dataset size will lead to coverage differences that affect polishing efficacy and therefore Q-score. These scores are indicative of performance but are not necessarily a controlled comparison.



A photograph of two men working on a laptop outdoors at night. The man on the right, wearing a blue hoodie with the number '9' on it, is holding a small device with a bright light illuminating his hands. The man on the left, wearing a dark jacket, is looking down at the laptop screen. A glowing keyboard is visible on the laptop. In the foreground, a smartphone displays a map. The background is dark, suggesting a nighttime outdoor setting.

@aphilosof

porecamp.github.io

78 degrees
North

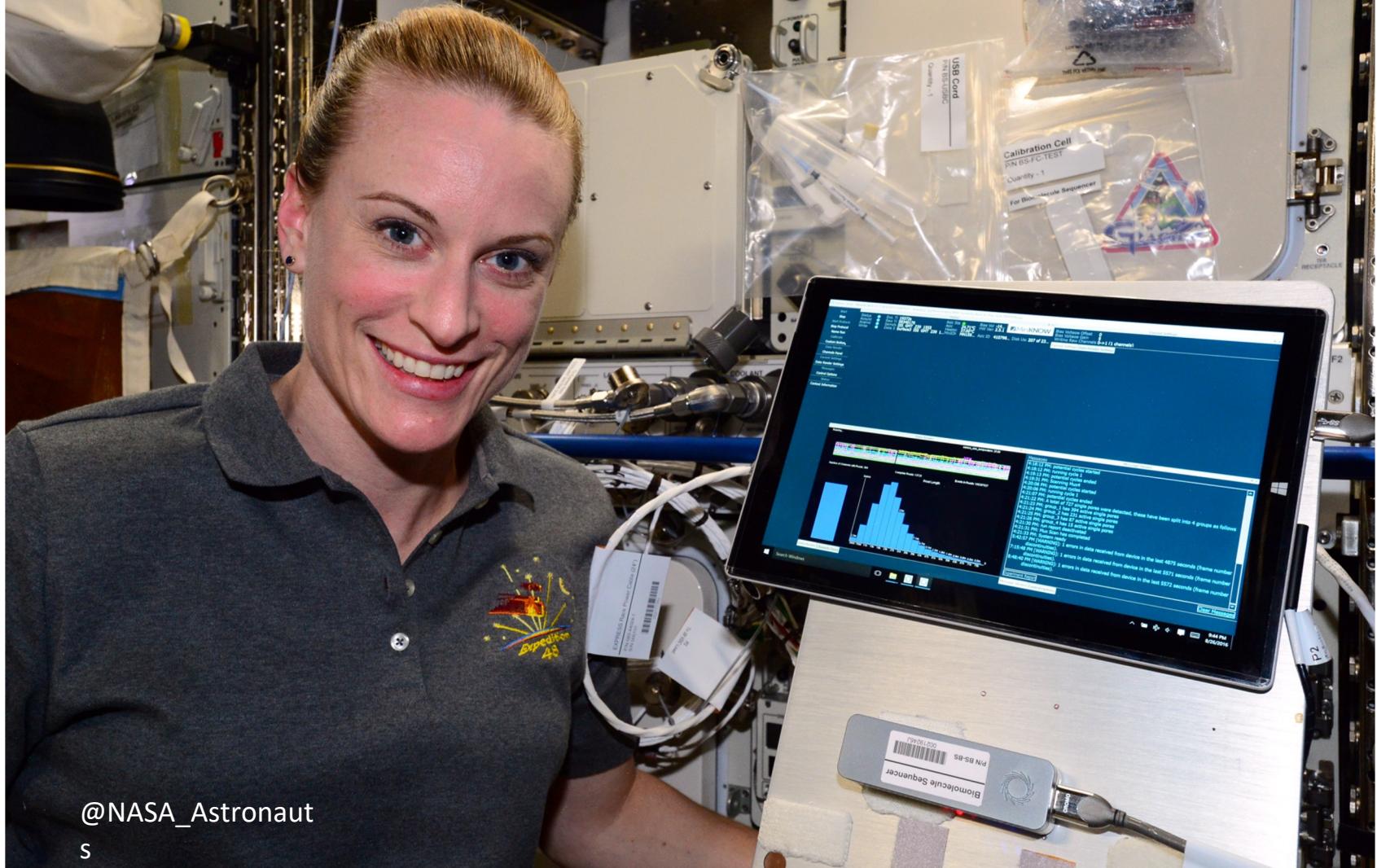


@arywnedwards





@explornau
t



@NASA_Astronaut

S

Ultra-long read methods

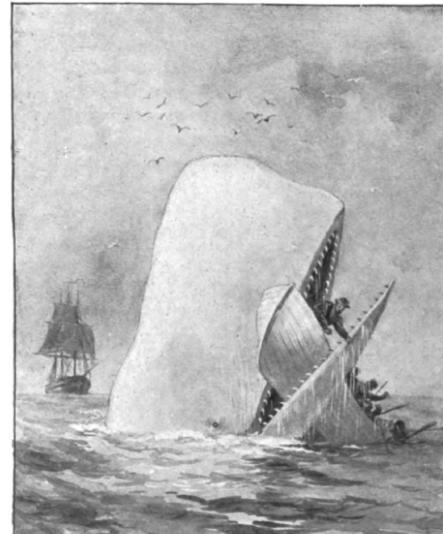
Thar she blows! Ultra long read method for nanopore sequencing

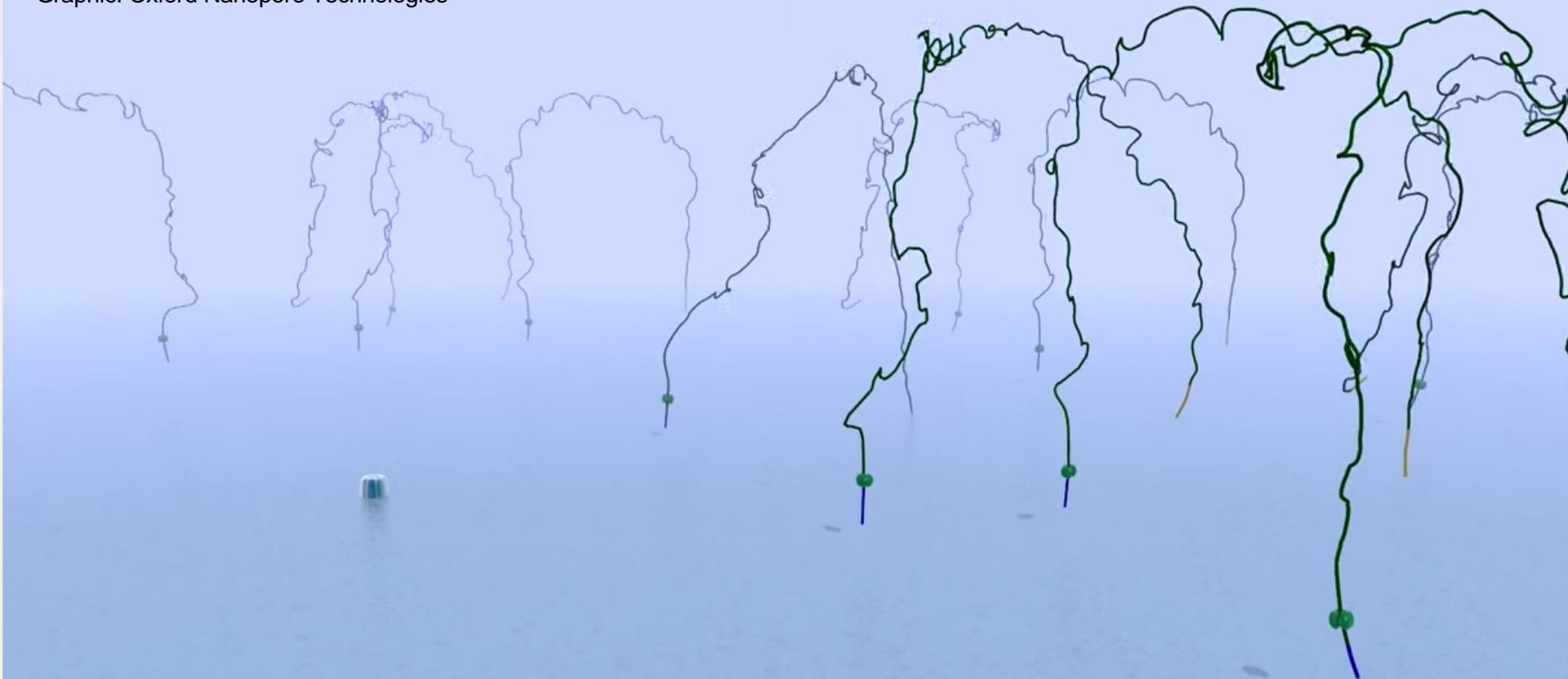
09 Mar 2017

tl;dr version

- Ultra long reads (up to 882 kb and indeed higher) can be achieved on the Oxford Nanopore MinION using traditional DNA extraction techniques and minor changes to the library preparation protocol, without the need for size selection
- The [protocol is available here](#); it involves a modified Sambrook phenol-chloroform extraction/purification, DNA QC, minimal pipetting steps, high-input to rapid kit and MinKNOW 1.4
- We have tested it on *E. coli* and human so far with good results; data is of course available

Ultra-long reads: background

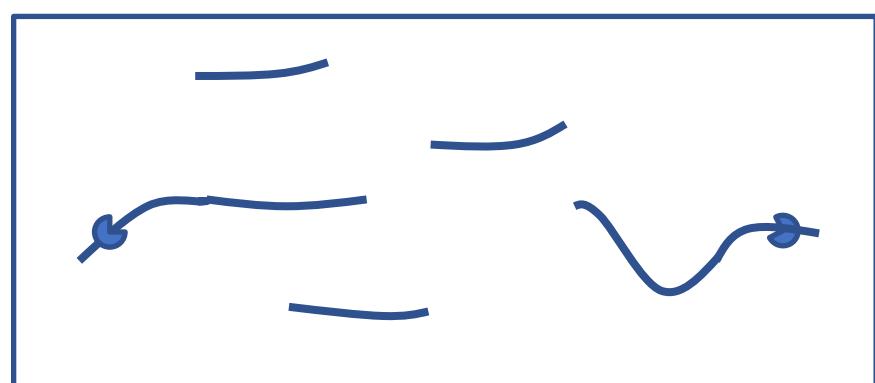
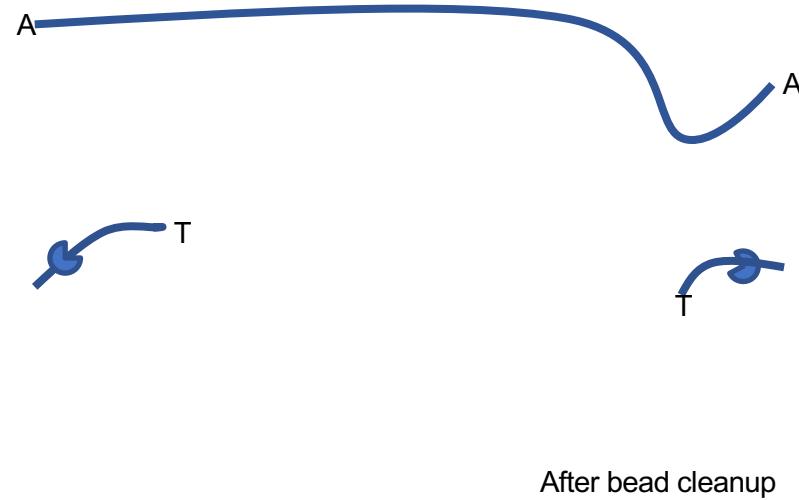




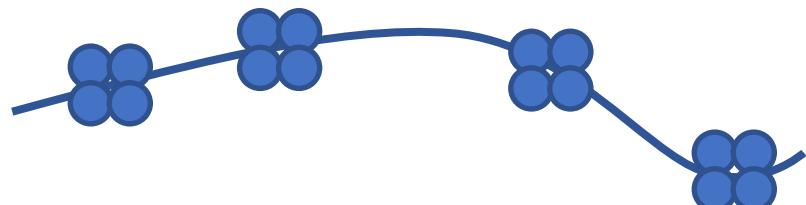
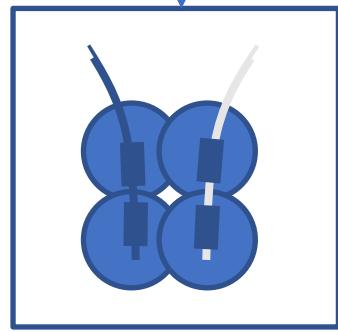
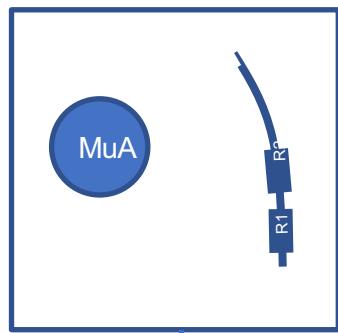
Flowcell yield dependent on input molarity

Ligation libraries

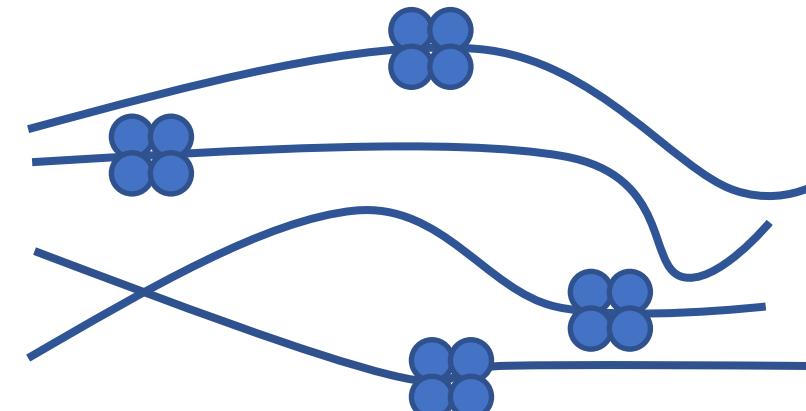
- Fragments sheared during preparation
- Rapid libraries fewer steps and no clean-ups



Transposase libraries for ultra-long reads

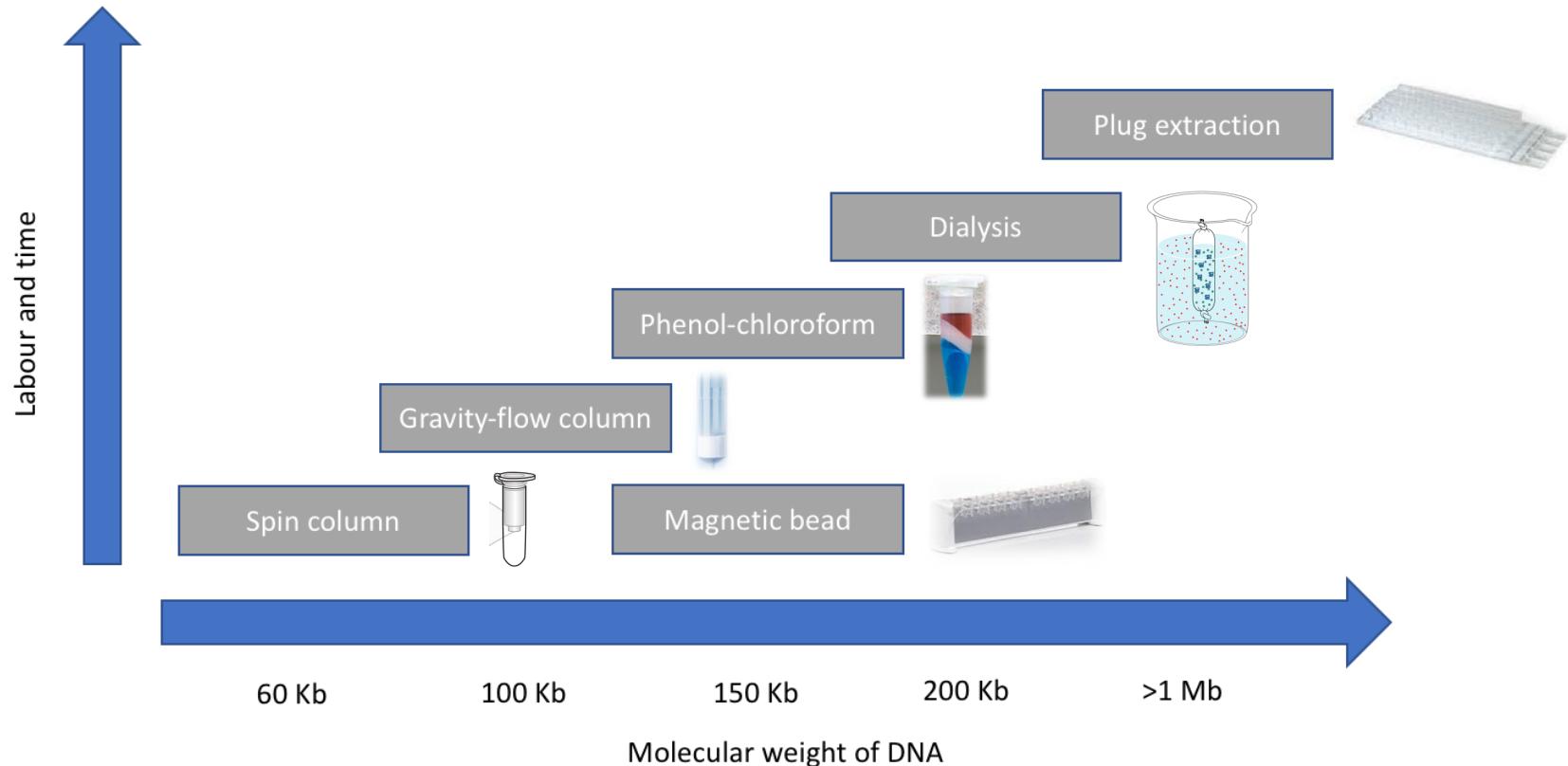


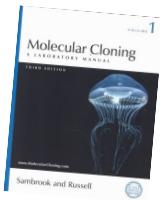
Standard
rapid



Ultra-long

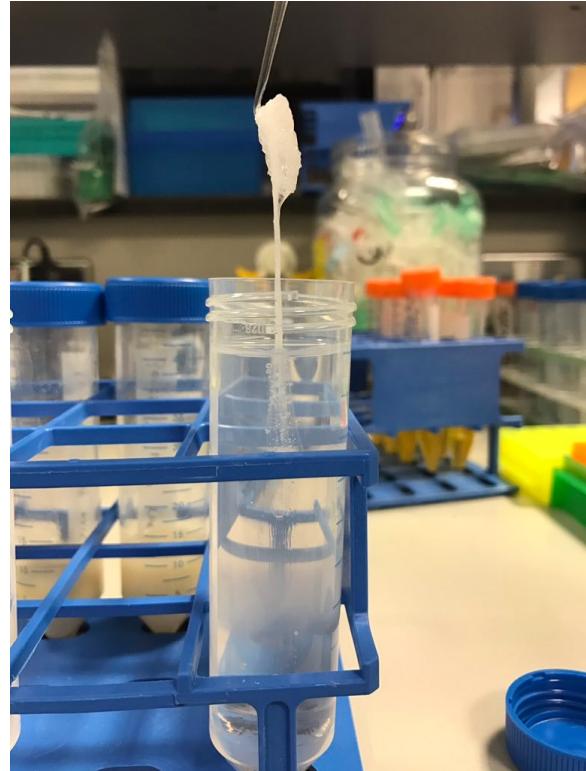
DNA preparation is critical





The Sambrook method

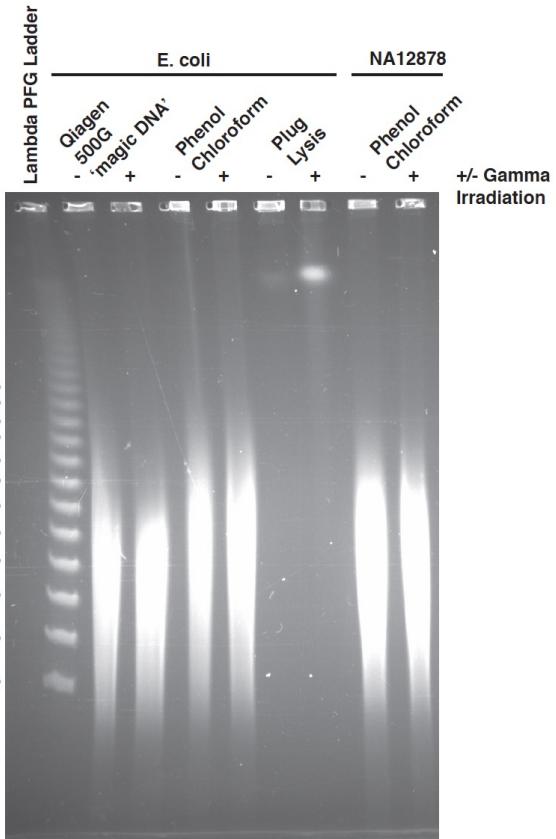
Step	Method
Lysis	0.5% SDS
Digestion	Proteinase K
Deproteinization	Phenol/Chloroform
Ethanol precipitation	Salt and ethanol
Resuspension	Elution buffer



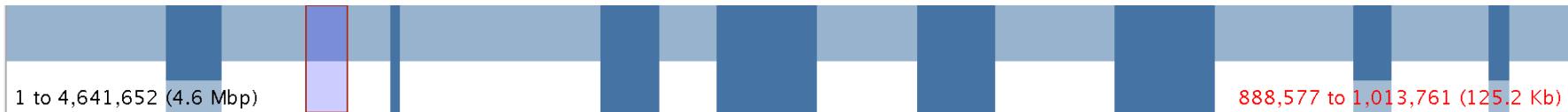
E. coli dataset

- High input phenol/chloroform library (250 ng/ul)
- Theoretical coverage 1x in reads >500 Kb

	<i>E. coli</i> run
Yield	5 Gb
Mean	33.3 Kb
N50	63.7 Kb
Max	778 Kb



E. coli assembly in 8 reads



Read	Length	Ref start	Ref end	Time (m)
1	876991	4398844	634183	32.48
2	696402	470003	1166405	25.79
3	799047	1137438	1936485	29.59
4	642071	1759431	2401502	23.78
5	826662	2106227	2932889	30.61
6	883962	2699626	3583588	32.73
7	825191	3285196	4110387	30.56
8	463341	3995967	4459308	17.16

miniasm assembly
N50: 4 Mb
Time: 1.5 s (1 CPU)

Nanopore human genome consortium

- Data:

<https://github.com/nanopore-wgs-consortium/NA12878>

- Protocol:

<https://dx.doi.org/10.17504/protocols.io.mrxc57n>

Birmingham



East Anglia



Nottingham



British Columbia



Santa Cruz



Springer Nature. All rights reserved.

nature
biotechnology
OPEN

Nanopore sequencing and assembly of a human genome with ultra-long reads

Miten Jain^{1,13} , Sergey Koren^{2,13}, Karen H Miga^{1,13}, Josh Quick^{3,13}, Arthur C Rand^{1,13}, Thomas A Sasani^{4,5,13} , John R Tyson^{6,13}, Andrew D Beggs⁷ , Alexander T Dilthey⁷ , Ian T Fiddes¹, Sunir Malla⁸, Hannah Marriott⁸, Tom Nieto⁷, Justin O'Grady⁹ , Hugh E Olsen¹, Brent S Pedersen^{4,5}, Arang Rhee² , Hollian Richardson⁹, Aaron R Quinlan^{4,5,10} , Terrance P Snutch⁶, Louise Tee⁷, Benedict Paten¹, Adam M Phillippy², Jared T Simpson^{11,12}, Nicholas J Loman³ and Matthew Loose⁸

We report the sequencing and assembly of a reference genome for the human GM12878 Utah/Cephal cell line using the MinION (Oxford Nanopore Technologies) nanopore sequencer. 91.2 Gb of sequence data, representing ~30x theoretical coverage, were produced. Reference-based alignment enabled detection of large structural variants and epigenetic modifications. *De novo* assembly of nanopore reads alone yielded a contiguous assembly (NG50 ~3 Mb). We developed a protocol to generate ultra-long reads (N50 > 100 kb, read lengths up to 882 kb). Incorporating an additional 5x coverage of these ultra-long reads more than doubled the assembly contiguity (NG50 ~6.4 Mb). The final assembled genome was 2,867 million bases in size, covering 85.8% of the reference. Assembly accuracy, after incorporating complementary short-read sequencing data, exceeded 99.8%. Ultra-long reads enabled assembly and phasing of the 4-Mb major histocompatibility complex (MHC) locus in its entirety, measurement of telomere repeat length, and closure of gaps in the reference human genome assembly GRCh38.

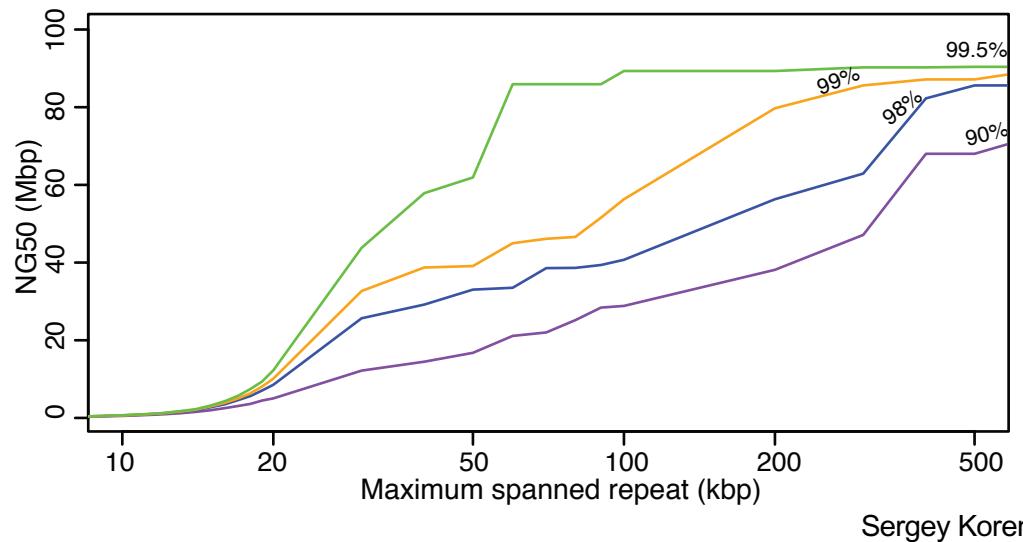
bioRxiv preprint doi: <https://doi.org/10.1101/2018.05.10.258006>; this version posted May 10, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

[dx.doi.org/10.1038/nbt.4060](https://doi.org/10.1038/nbt.4060)

ARTICLES

Why ultra-long reads?

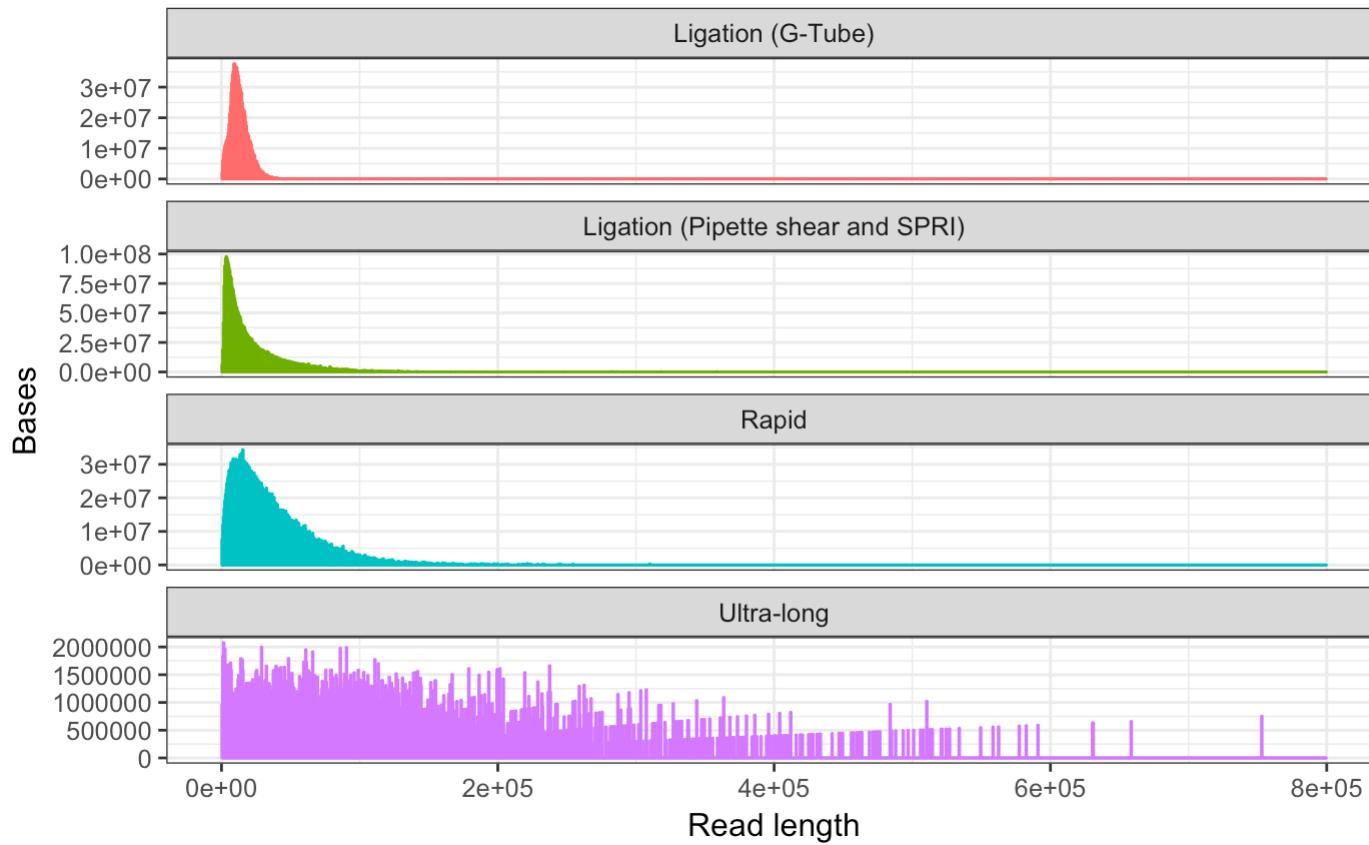
“Most recently, Nanopore sequencing reads approaching 1 Mbp were reported and it is imaginable that further technology advances will enable the complete assembly of a diploid human within a few years.”



Adam Phillippy

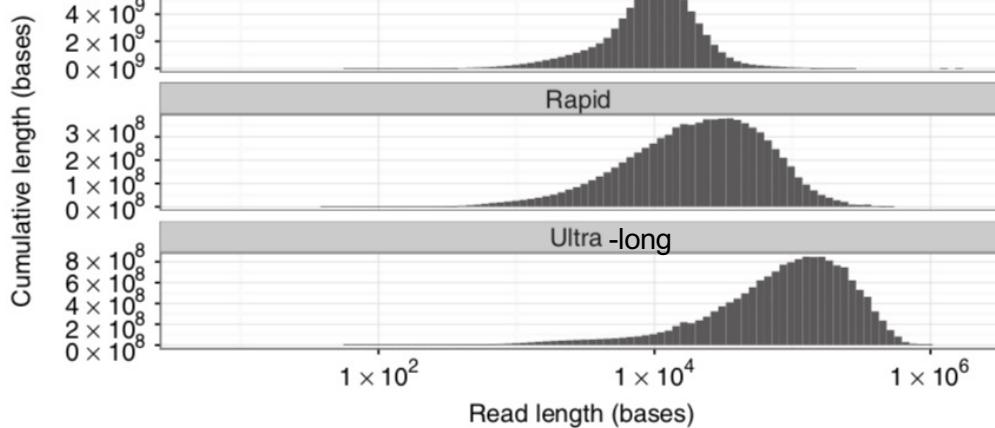
Sergey Koren

Ultra-long reads



Comparison of read sets

b



Runs	Total (Gb)	Yield mean(Gb)	N50 (Kb)
39	91.2	2.3	10.5
4	8.5	2.1	30.4
12	16.8	1.4	99.8

Updated assembly

	Ligation 10kb assembly rel3	Rel3 + 5x ultras	20x regular + 10x ultras
Coverage (X)	~30	~35	~30
Miniasm N50 (Mb)	4.3	6.1	10.2
Largest contig (Mb)	17 Mb	32.29Mb	43.3 Mb
Params	minimap1	minimap1	minimap2 ava-ont -r 10000
Canu NG50 (Mb)	3.0	6.4	n/a

Thanks to Heng Li for minimap2 parameters

The whale scale

Whale > Mass	
Killer whale: 3,600 – 5,400 kg	
Blue whale: 140,000 kg	
Humpback whale: 30,000 kg	
Beluga whale: 1,400 kg	
Sperm whale: 35,000 – 57,000 kg	
North Pacific right whale: 50,000 – 8...	
Narwhal: 940 kg	
North Atlantic right whale: 40,000 – 7...	
Short-finned pilot whale: 1,000 – 3,0...	

1 kb == 1 kg

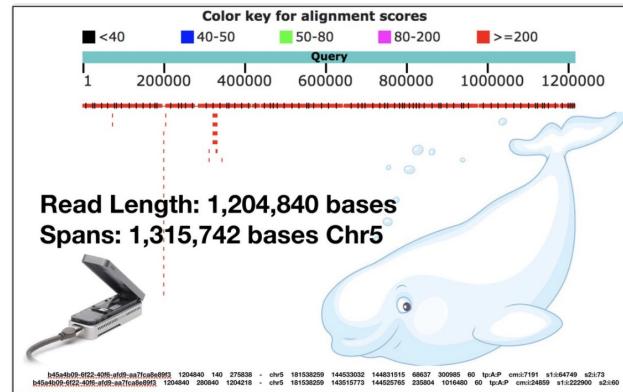


Matt Loose

@mattloose

Follow

Amazing what you can see on an evenings whale watching with [@nanopore](#) - I saw a big one this tonight! Thanks to Nadine, [@alexomics](#) and of course [@scalene](#) . Protocol tweaks will be available shortly on lab.loman.net/protocols



Current ultra-long read methods



@longreadclub
youtube.com/longreadclub

John Tyson
@DrT1973

Following

Right then.... Posted our methods for HMW extraction, ligation library protocol modifications and bead-free methods for increasing 100Kb+ ultra-long reads on the @nanopore community.

community.nanoporetech.com/posts/rocky-mo...

Will also be pushing to @longreadclub and @protocolsIO soon enjoy 😊

Bead Free Long Fragment LSK109 Library Prep

```
graph TD
    A[Genomic DNA Fragmentation] --> B[End-Prep]
    B --> C[Add Equal volume (20 μl PEG8000 (w/v), 1M NaCl) Incubate 30mins, Spin 13K rpm 20mins, 2x 70% EtOH Wash Resuspend in EB]
    C --> D[Adapter Ligation]
    D --> E[Add NaCl to 400mM Incubate 20mins, Spin 13K rpm 30mins 2x Wash (4.0% PEG8000, 500mM NaCl) Resuspend in EB]
    E --> F[Load]
    F --> G[P1000 tip]
```

P1000 tip

21% seq in reads >100Kb

100kb

Estimated Read Length in Bases

1:18 PM - 9 May 2019

101 Retweets 217 Likes

8 101 217

New HMW DNA extraction kits



Ultra-long kit

Ultra-long sequencing kit (SQK-ULK114)

Based on the rapid kit

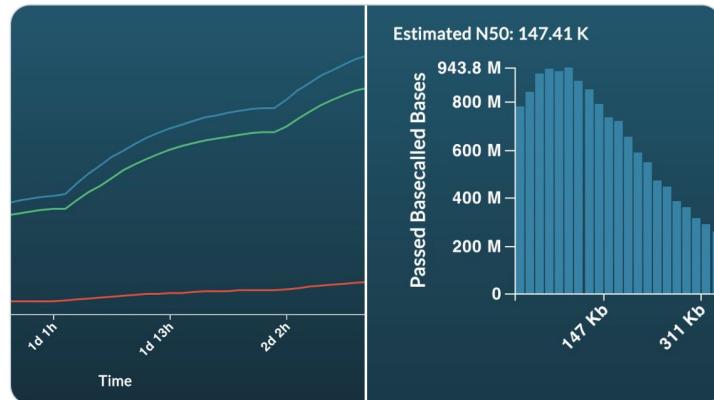
Combines fragmentation at lower concentration and clean-up after adapter attachment

These changes could be applied to other extractions



Matt Loose
@mattloose

I still remember the first time I saw data from [@scalene](#) showing ultra long reads ($N_{50} \geq 100\text{kb}$) were possible. Since then we and others (especially [@DrT1973](#)) have worked to get longer reads. This weekend we tested the [@circulomics](#) [@nanopore](#) ultra long kits. Wow! 19.49 Gb N_{50} 143kb



9:18 AM · Feb 1, 2021 · Twitter Web App

37 Retweets 5 Quote Tweets 126 Likes

Do you really need a kit?

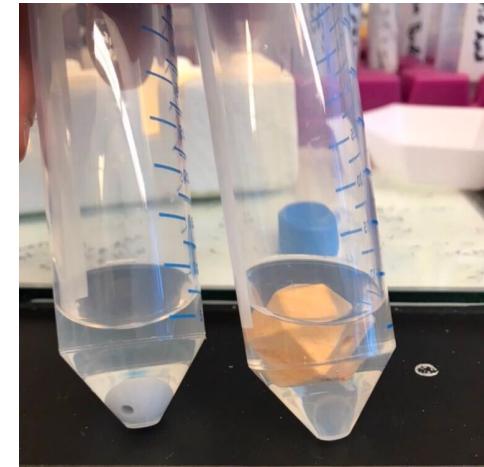
Standard SDS/ProtK lysis

Salt/PrOH precipitation

DNA will bind many things



???



@DrT1973

Acknowledgments

University of Birmingham: Nick Loman, Sam Nicholls, Radoslaw Poplawski

University of Vermont: Scott Tighe

University of British Columbia: John Tyson

Oxford Nanopore Technologies: Divya Mirrington

Zymo Research: Shuiquan Tang, Ryan Sasada, Ryan Kemp

Cambridge Biosciences: Hannah McDonnell

Motivation for long read metagenomics

Want to apply this new long read extraction technique and ideally have single contig *de novo* assembly of genomes, from complex samples

- improve power for taxonomic assignment
- phylogenetics on whole genomes
- identify strain-level variation (haplotyping)

For application to clinical translational projects in Birmingham

- evaluating faecal microbiome transplants for ulcerative colitis (STOP-COLITIS)
- evolution of the respiratory microbiome in cystic fibrosis (with Uni. of Liverpool)

Long read metagenomics for faecal bacteriotherapy



UoBMTC

@UOBMicrobiomeTC

FMT FRIDAY! Always a pleasure seeing Lee @MidFreewheelers and John @Staffsbloodbike, laughter is guaranteed. Another 4 FMT treatment sets heading north and south today. Thank you all. #FMT #CDI #bloodbikes @unibirm_MDS



4:26 AM · Nov 8, 2019 · Twitter for iPhone

@UOBMicrobiomeTC

UOBMTC Retweeted



John Gordon @John_N_Gordon · Feb 12

Our first @HHFT colonoscopy delivered #FMT (faecal microbial transplant) today being administered by @DavidAJLloyd. Really pleased we've got this up and running again and hoping it will make a big difference to our patients with #CDAD and #IBD.



2

2

17



Benchmarking with mock community



(GIGA)ⁿ
SCIENCE

GigaScience, 8, 2019, 1–9

doi: 10.1093/gigascience/giz043
Data Note

DATA NOTE

Ultra-deep, long-read nanopore sequencing of mock microbial community standards

Samuel M. Nicholls ^{1,†}, Joshua C. Quick ^{1,†}, Shuiquan Tang ² and Nicholas J. Loman ^{1,*}

¹Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, Edgbaston, B15 2TT, UK; and ²Zymo Research Corporation, 17062 Murphy Ave., Irvine, CA 92614, USA

*Correspondence address. Nicholas J. Loman, Institute of Microbiology and Infection, School of Biosciences, University of Birmingham, UK. E-mail: n.j.loman@bham.ac.uk <http://orcid.org/0000-0002-9843-8988>

[†]Contributed equally.

PromethION

Exciting opportunity for higher yields

\$2000/flowcell based on 12 pack

Twin 10 Gbps fibre

Twin Tesla V100 GPU cards (basecalling)

3,000 channels v 512 on MinION (6x)

Up to 24 flowcells in parallel

Paul Loman
@donlom

Following

Cheers @pathogenomenick and @Scalene for our lab tour and a chance to see your impressive new bit of kit



1:01 AM - 28 Jul 2018

2 Retweets 18 Likes

2 18

Mock community

ZymoBIOMICS Microbial Community Standard

Even or log distributed (10² to 10⁸ cells)

2 ug/prep (even), 220 ng/prep (log)

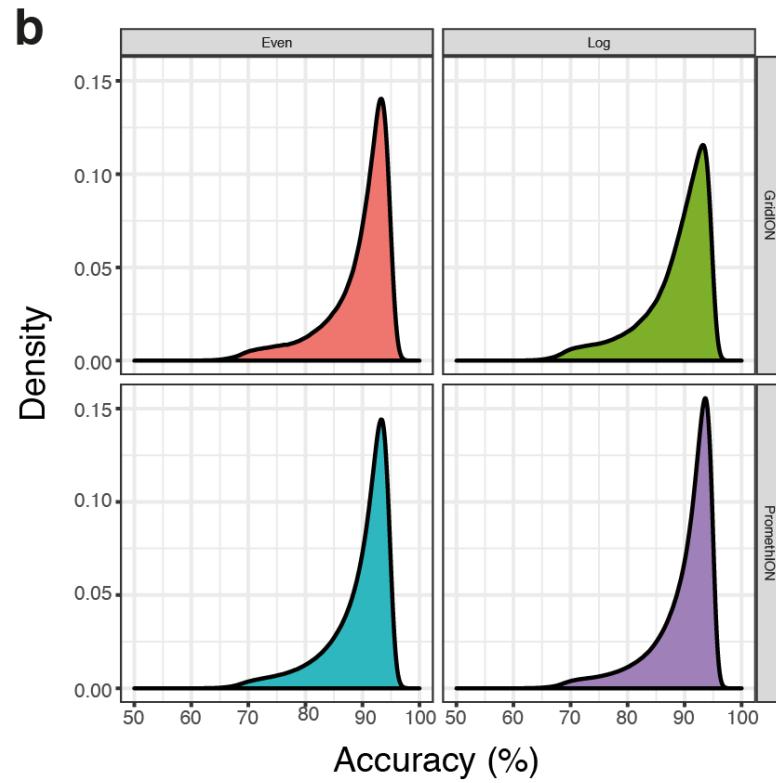
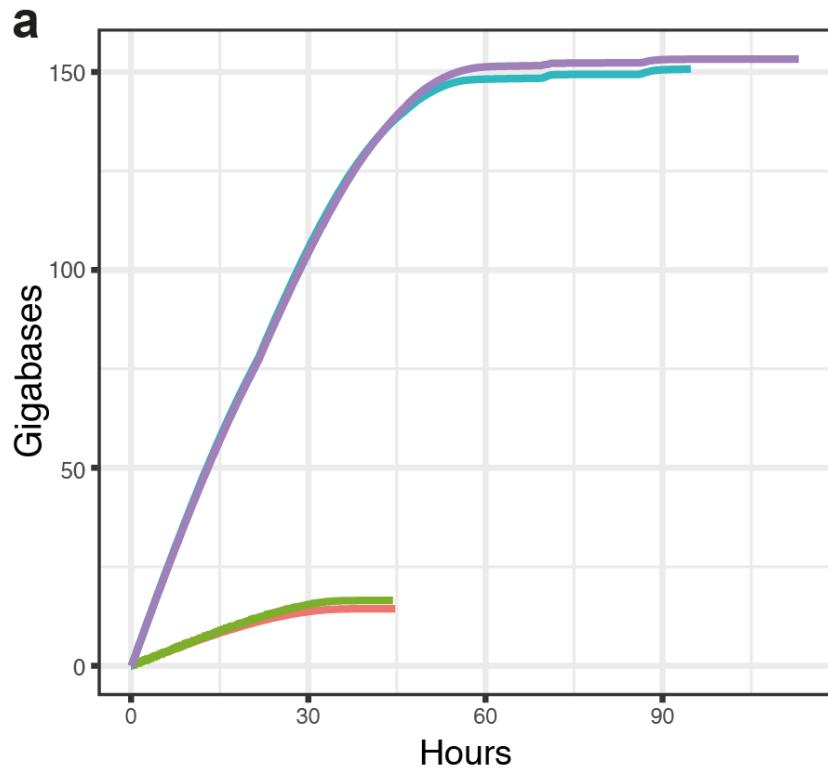
3 Gram-negative, 5 Gram-positive, 2 Fungi

P. aeruginosa, E. coli, S. enterica, L. fermentum, E. faecalis, S. aureus, L. monocytogenes, B. subtilis, S. cerevisiae, C. neoformans

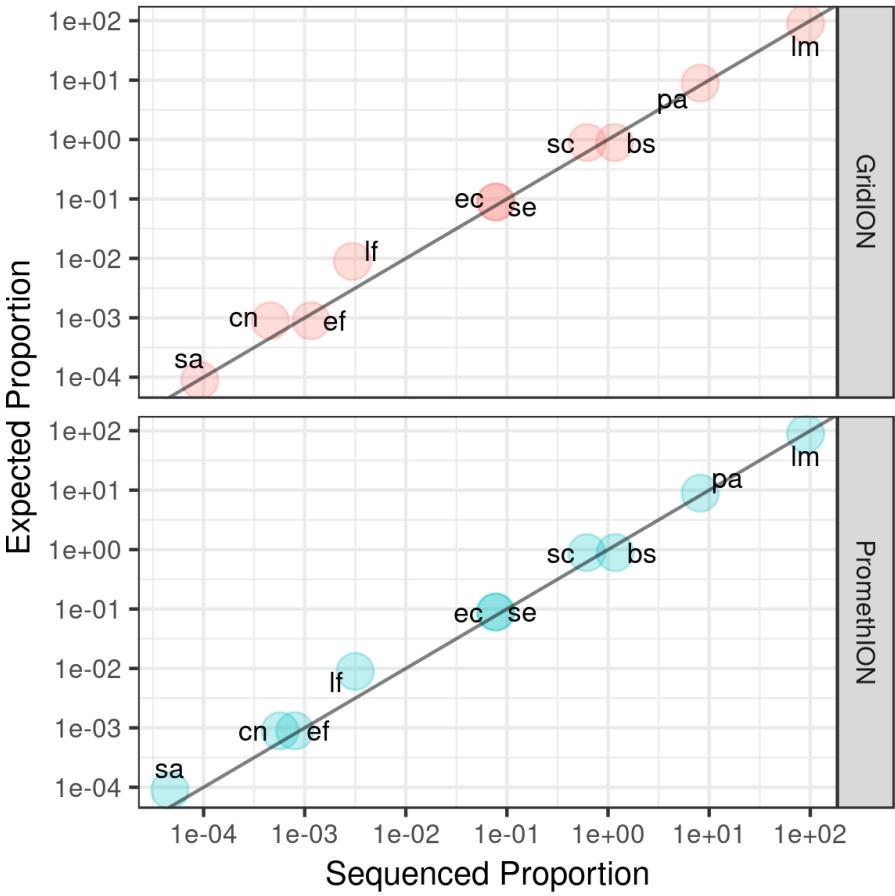
'Cellular' - but shipped in DNA/RNA Shield (lysis buffer)

Illumina data available from Zymo

Run statistics



Alignment stats



Sequencing mock community (Log)

Species	GridION			PromethION		
	Yield (Gbp)	Aln. N50 (Kbp)	Coverage (×)	Yield (Gbp)	Aln. N50 (Kbp)	Coverage (×)
<i>Listeria monocytogenes</i>	12.10	4.95	4043.90	110.09	4.97	36 796.21
<i>Pseudomonas aeruginosa</i>	1.10	9.38	161.45	9.99	9.33	1471.41
<i>Bacillus subtilis</i>	0.16	5.03	38.67	1.44	5.04	356.00
<i>Saccharomyces cerevisiae</i>	0.08	4.78	6.93	0.75	4.75	62.33
<i>Salmonella enterica</i>	0.01	9.20	2.20	0.10	9.17	20.04
<i>Escherichia coli</i>	0.01	8.65	2.14	0.09	9.17	19.24
<i>Lactobacillus fermentum</i>	4×10^{-4}	3.40	0.210	0.004	3.37	2.03
<i>Enterococcus faecalis</i>	2×10^{-4}	7.62	0.055	1×10^{-3}	6.05	0.34
<i>Cryptococcus neoformans</i>	6×10^{-5}	4.41	0.003	7×10^{-4}	4.97	0.037
<i>Staphylococcus aureus</i>	1×10^{-5}	7.12	0.005	5×10^{-5}	3.58	0.020

Limit of detection

Total input 7.5×10^8 cells with lowest abundance organism *S. aureus* ~400 cells (~1 pg)

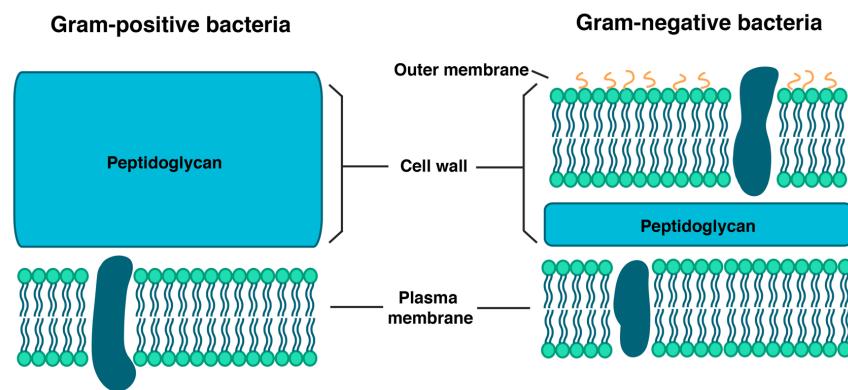
All organisms detected:

6/10 organisms high enough coverage to assemble

1/10 organisms could give gene presence/absence information

Recalcitrant organisms

- Gram-positive bacteria have a tough cell wall, both Gram-positive and negative can also have a polysaccharide or glycoprotein capsule
- Yeast may also have a capsule and have β -glucan layer, chitin layer and plasma membrane
- Spore forming bacteria have a spore coat (keratin), outer membrane, spore cortex (peptidoglycan, dipicolinic acid, calcium), spore wall (peptidoglycan) and plasma membrane



MetaPolzyme

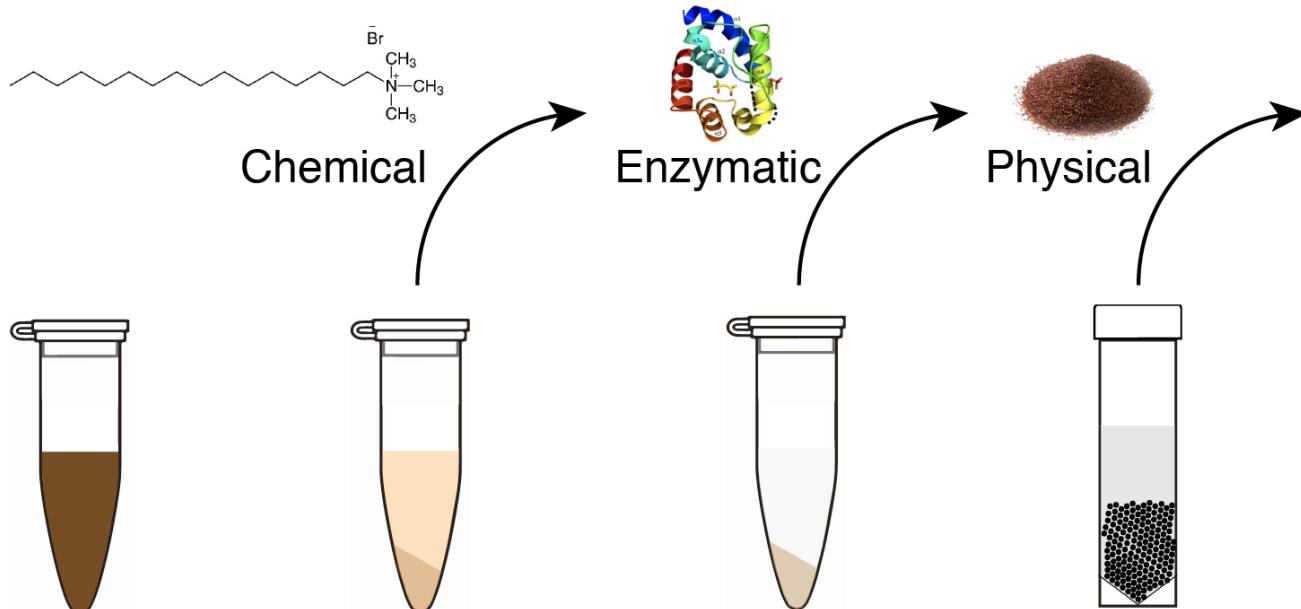
Originally developed by Scott Tighe for the ABRF Metagenomics Research Group (MGRG)

Cocktail of six enzymes for DNA extraction from metagenomic samples: mutanolysin, achromopeptidase, lyticase, chitinase, lysostaphin and lysozyme

Will generate spheroplasts/protoplasts in 1x PBS pH 7.5

Incubation time ~2 hours

3 peaks extraction



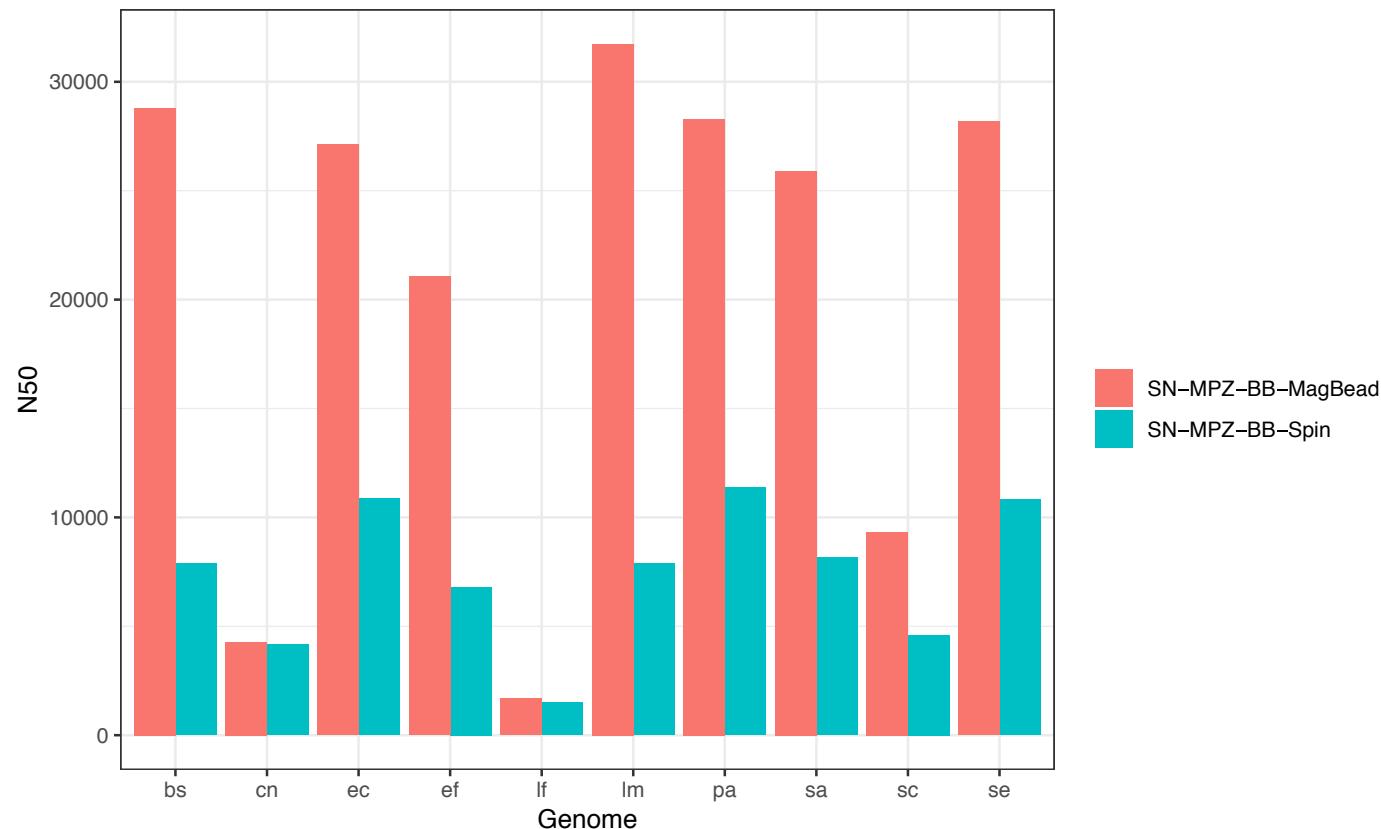
Sample in
OMNIgene GUT
or DNA/RNA Shield

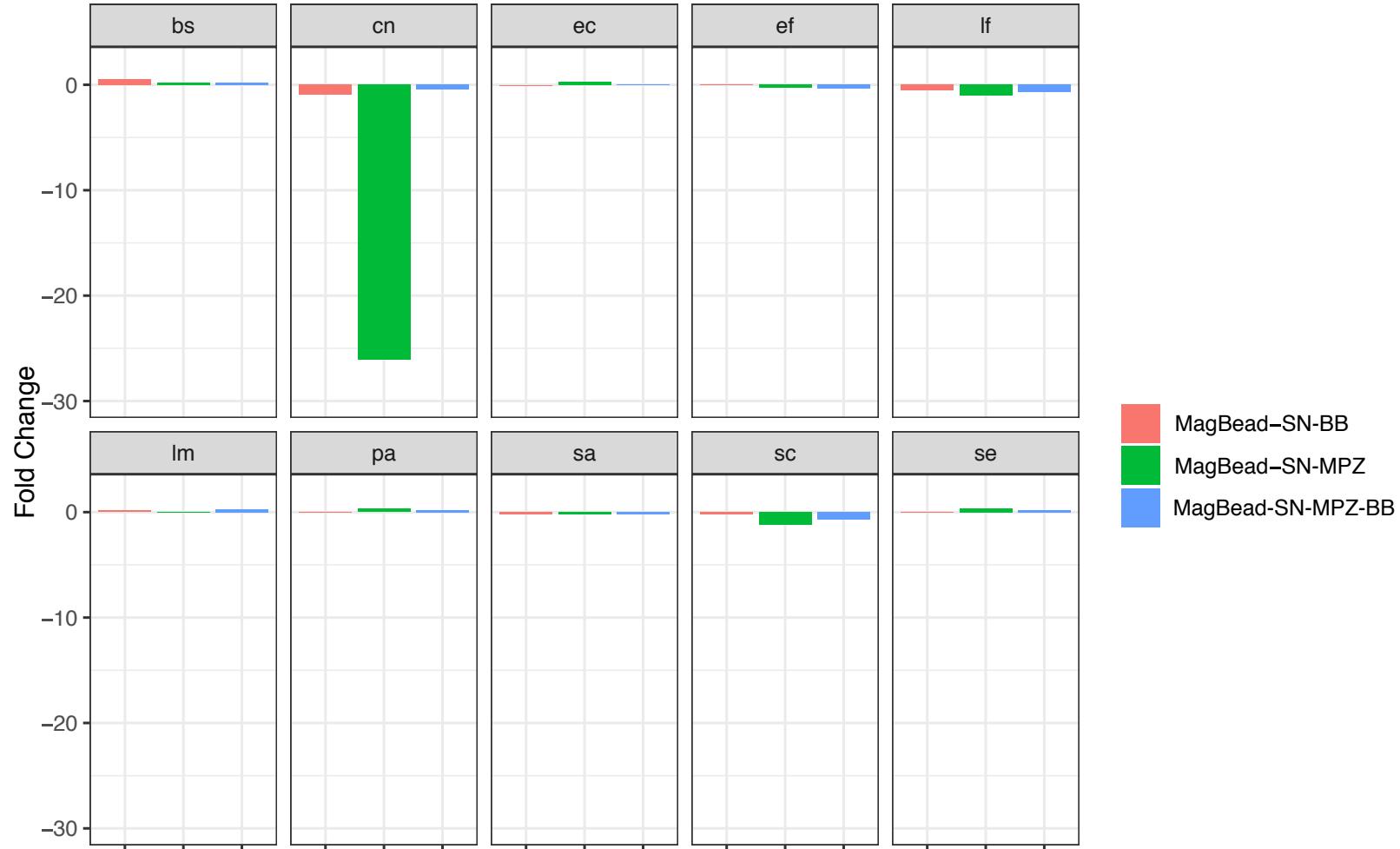
Supernatant
(SN)

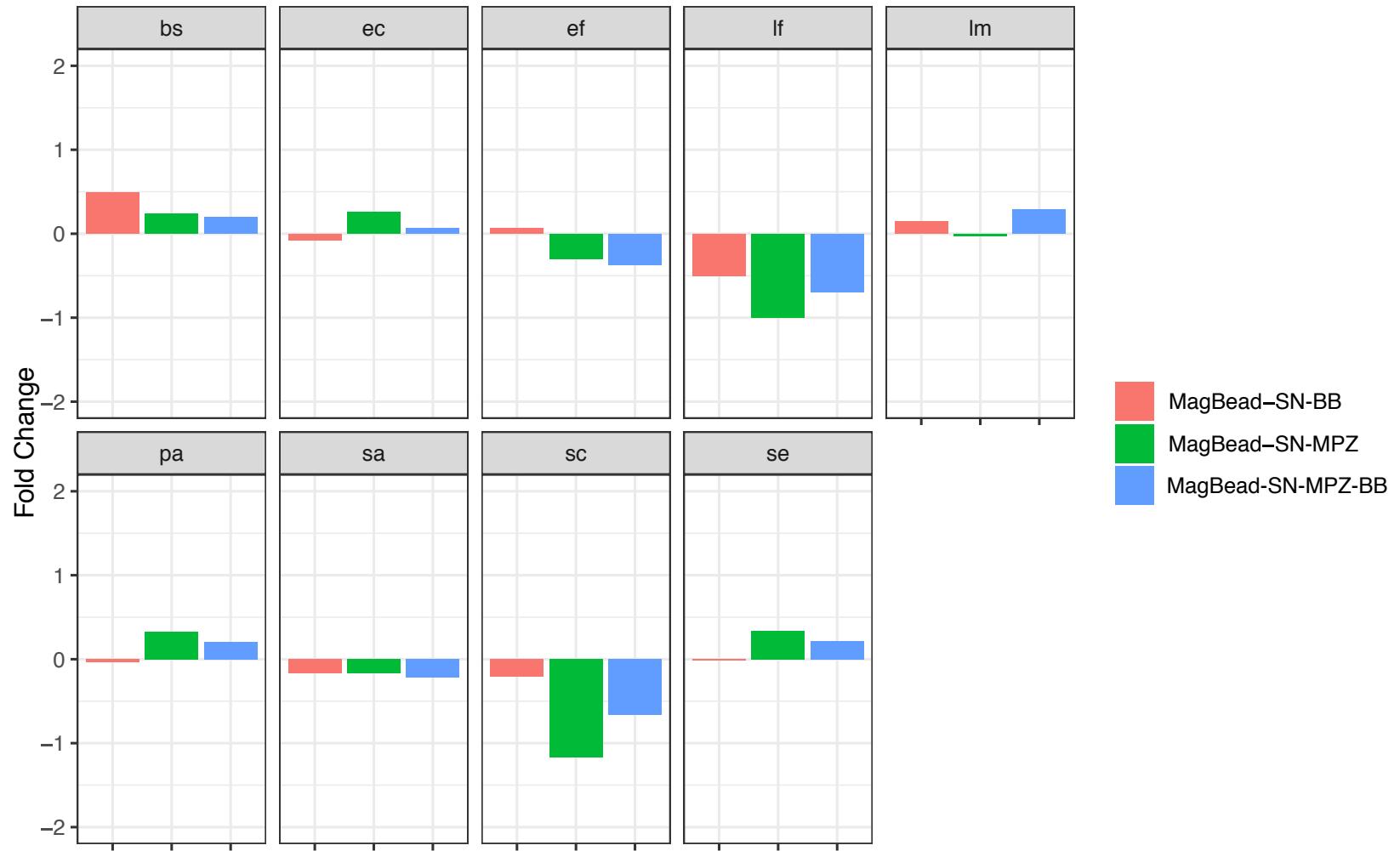
MPZ treatment
(MPZ)

Bead-beating
(BB)

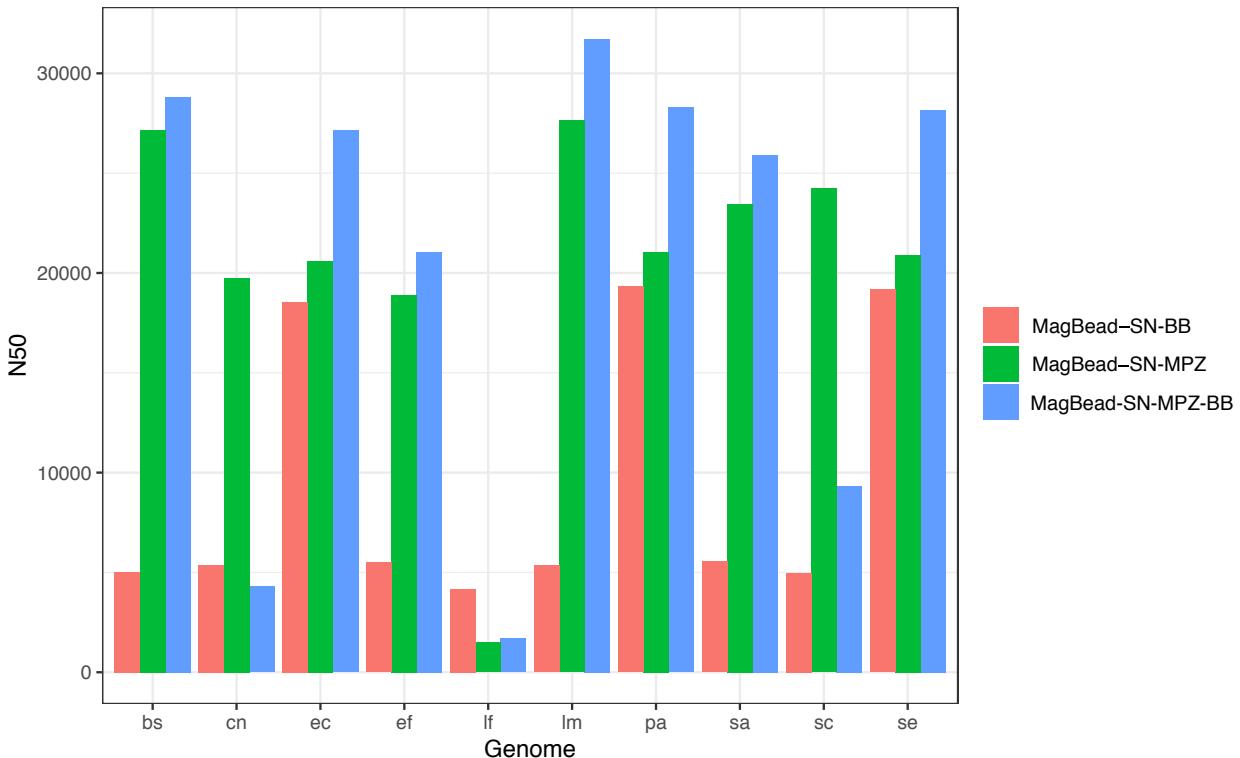
Spin columns do not support long-reads





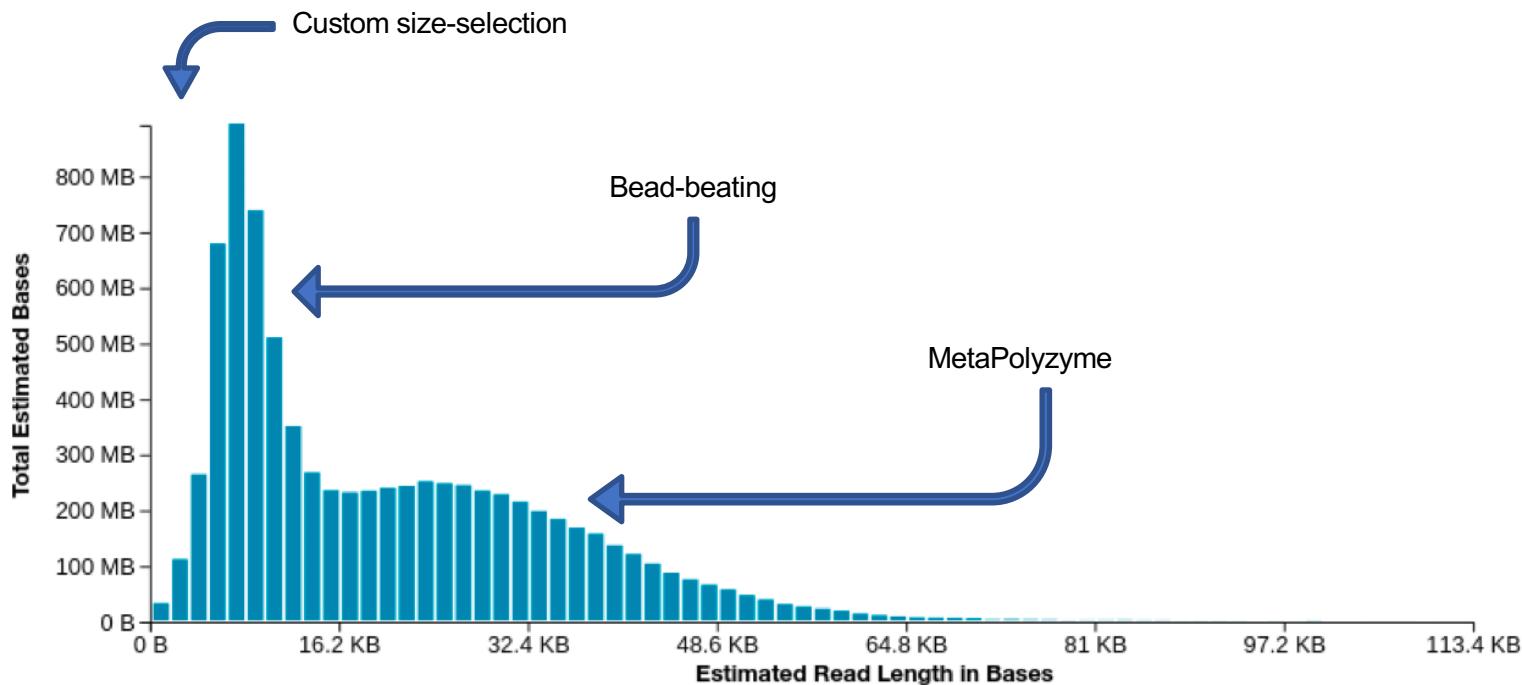


Optimised mock extraction



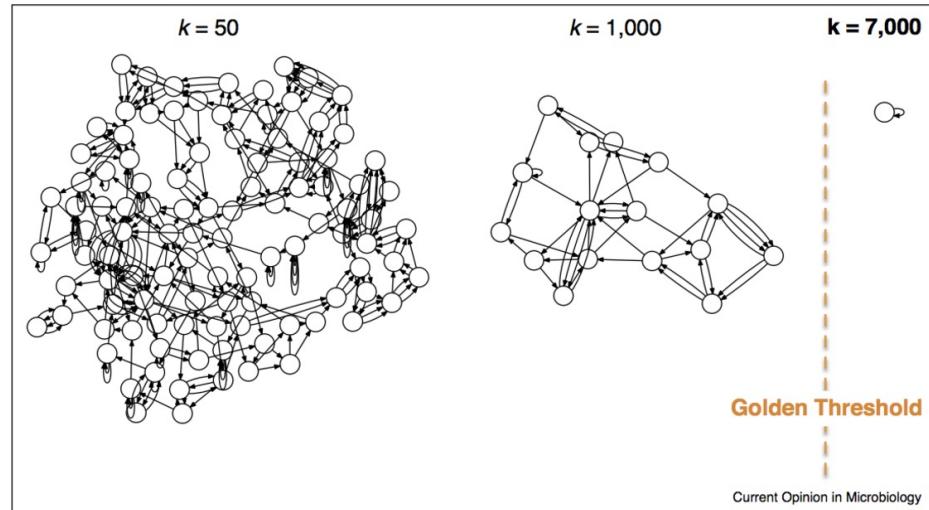
Optimised	
Yield (Gb)	8.1
Mean	8,124
N50	25,725

'3 peaks' length distribution



How long is long enough?

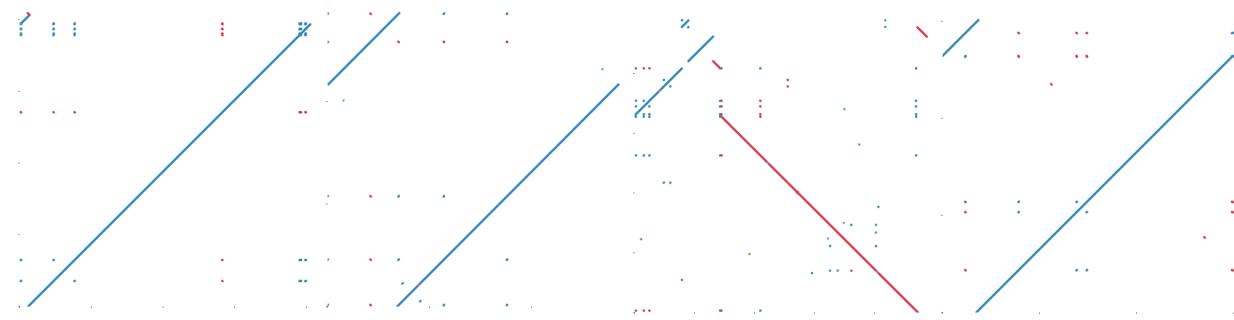
- 7kb reads will bridge rRNA operon
- Typically the longest bacterial repeat sequence



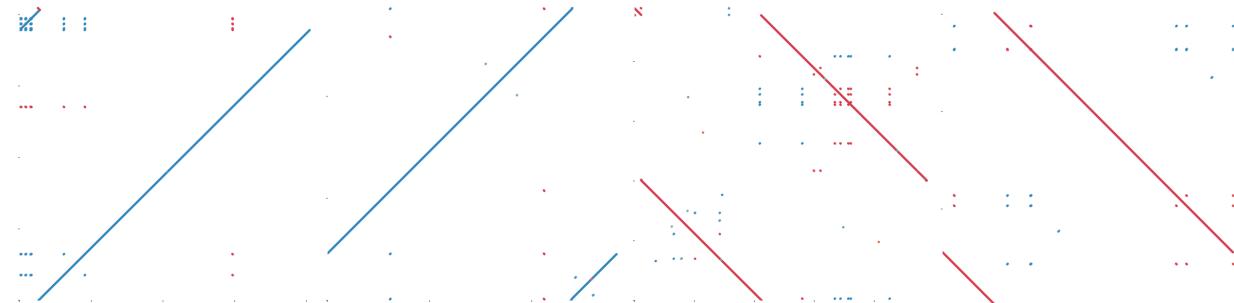
Read length simplifies the assembly problem. Simplified assembly graphs for the *Escherichia coli* K-12 genome are shown for varying read lengths k [26]. Contigs are nodes and unresolved adjacencies are represented as links. (a) With a short read length, many regions of the genome remain unresolved. (b) Increasing the length to 1 kbp, similar to that produced by Sanger sequencing [84], resolves a large fraction of the genome, but some complexity remains due to large global repeats. (c) Once the length is above the 'golden threshold,' which exceeds the most common repeat length in prokaryotic genomes, all ambiguity is removed from the graph (this figure was derived from the same data used in Figure 1 of [18**]).

Flye assemblies

Gigascience
dataset



Optimised



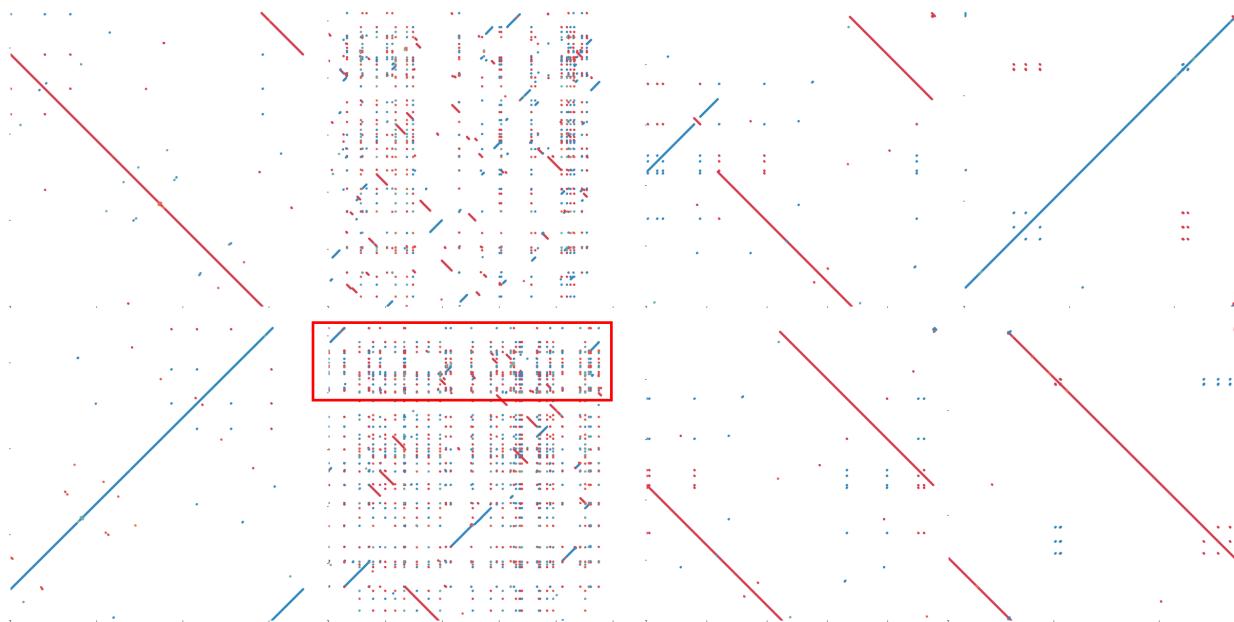
B. subtilis

E. faecalis

E. coli

L. monocytogenes

Flye assemblies cont.



Episode 4 - Flye with Mikhail Kolmogorov and Jeffrey Yuan
279 views • 4 weeks ago

Runs and yields on clinical samples

	FMT013_D1	FMT013_D10	FMT002_D1	FMT002_D56	FMT003_D1	FMT003_D1	FMT003_D56	FMT003_D56
Platform	PION	PION	MION	PION	MION	PION	MION	PION
Yield (Gb)	94.2	112.2	26.5	96.2	7.9	66.6	16.3	111.9
Mean	5,726	4,191	10,635	7,490	6,563	5,469	4,406	4,605
N50	7,950	6,356	16,015	8,358	16,713	10,343	6,857	7,572

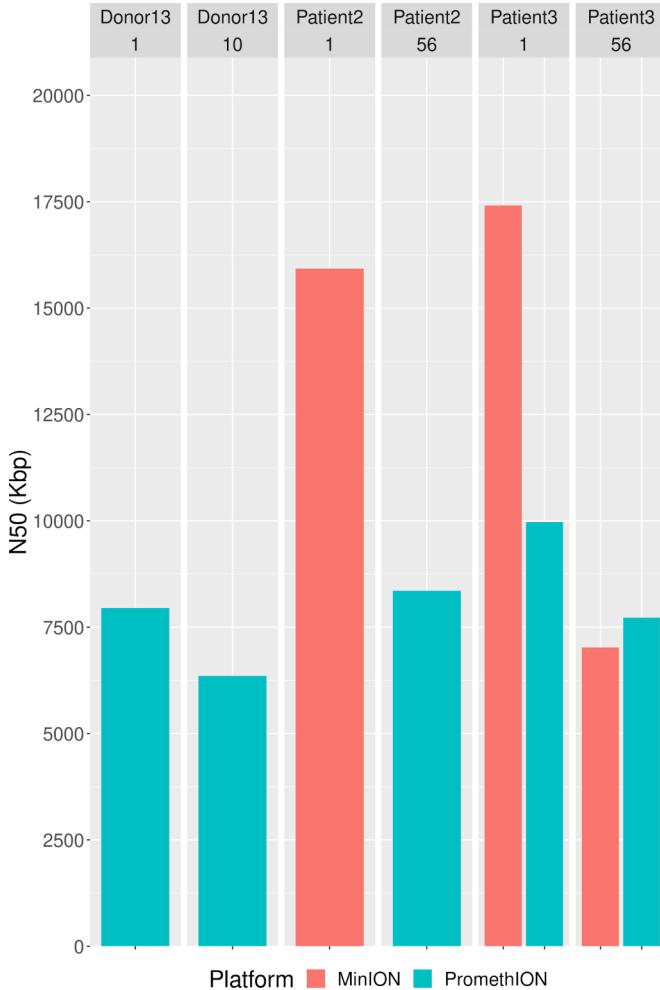
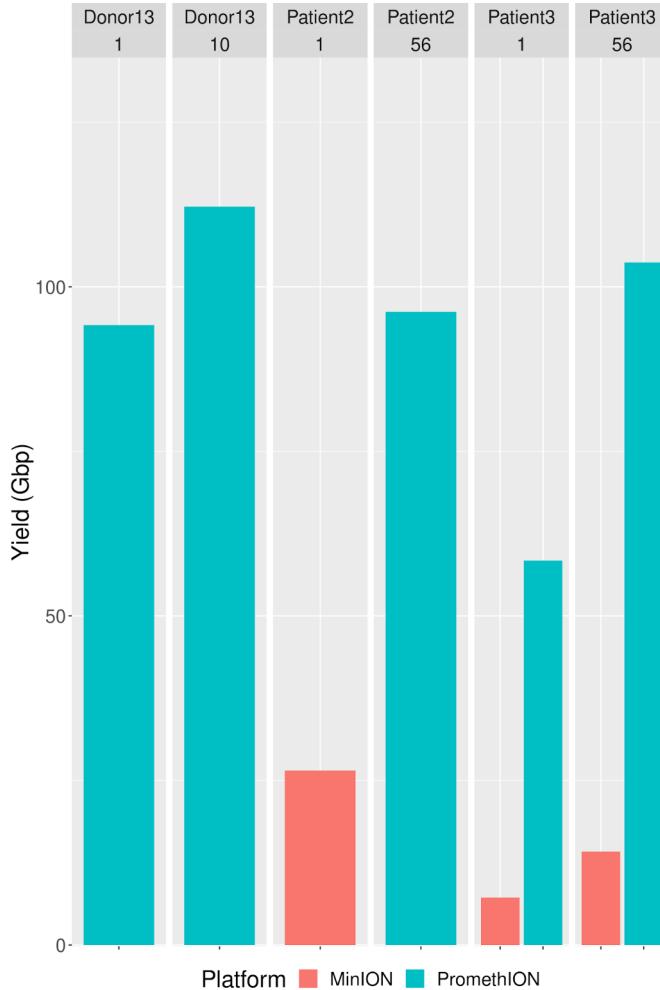
Long read FMT pilot

one donor

two recipients

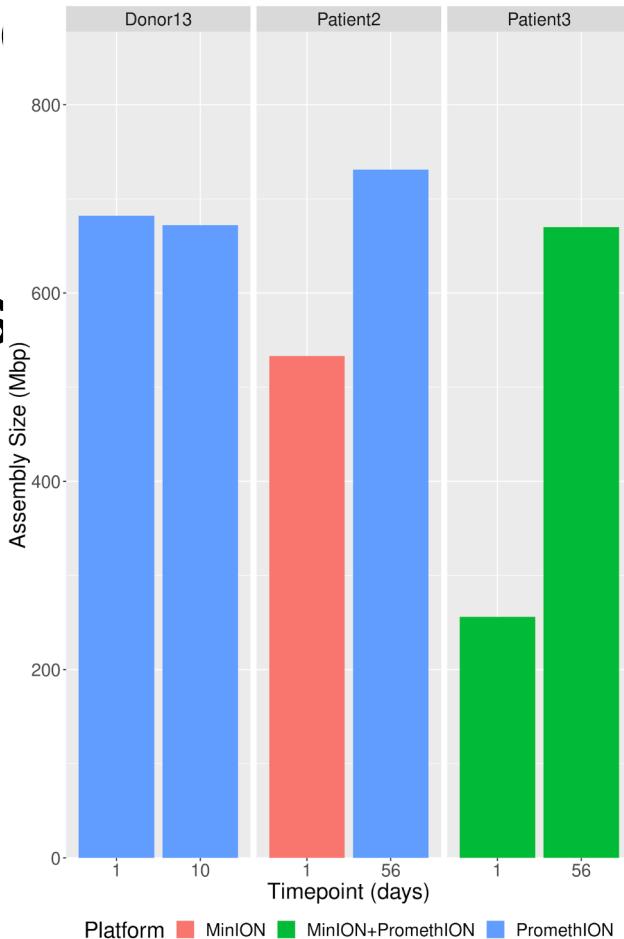
before / after

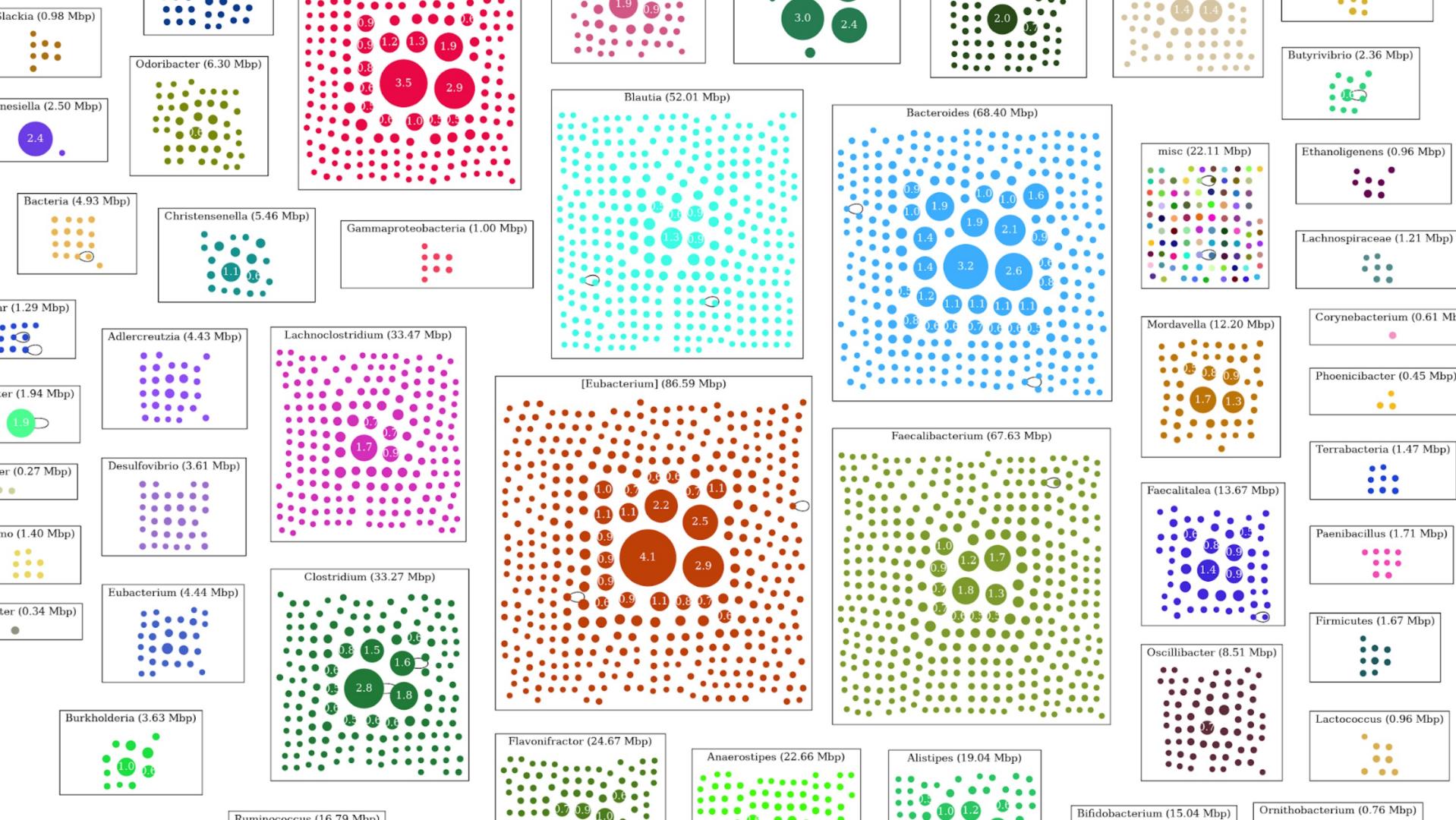
next stop
assembly





Remarkable contiguity from Flye assemblies





So, you have some big assemblies...

...but they are reflective of the underlying error rate of the reads
part of the reason flye and wtDBG2 are fast and efficient

assembly polishing is therefore a critical step towards high-quality
...but can be cumbersome and costly in terms of time and resources

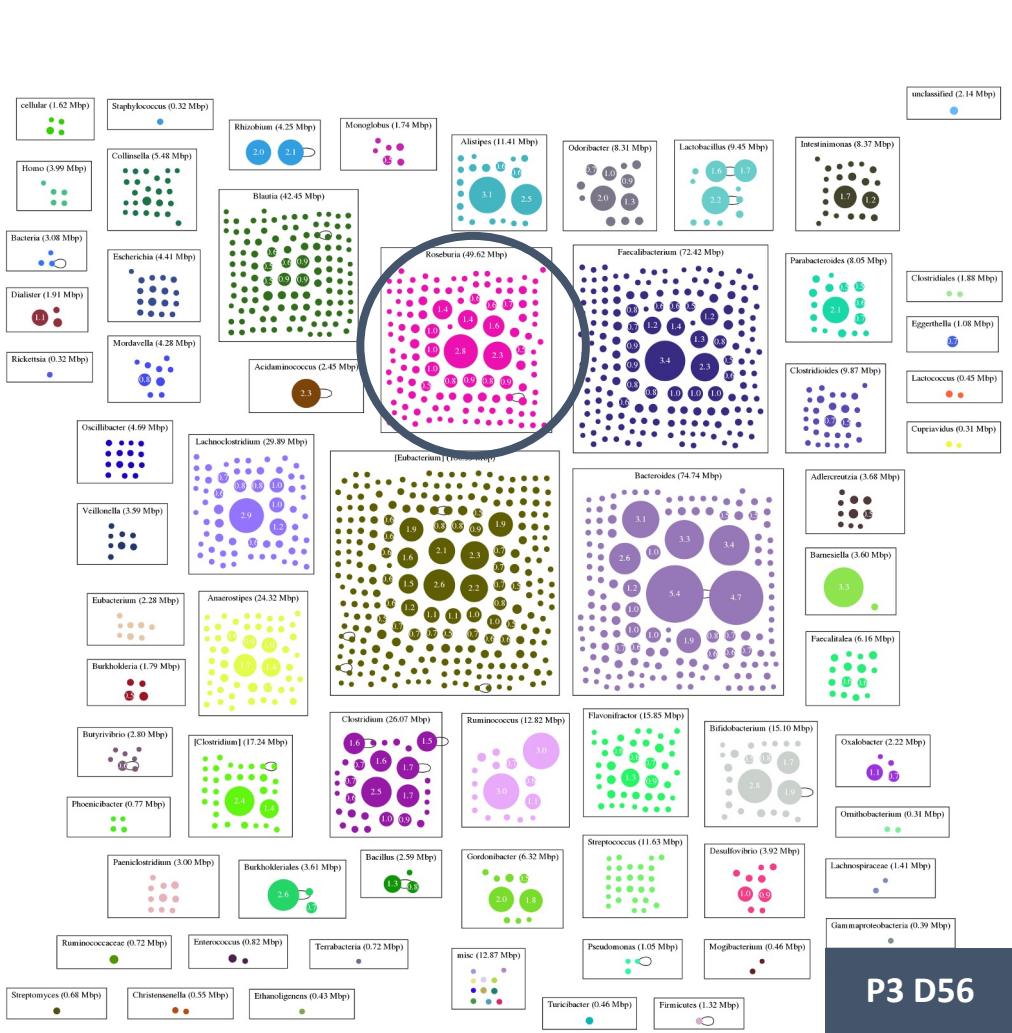
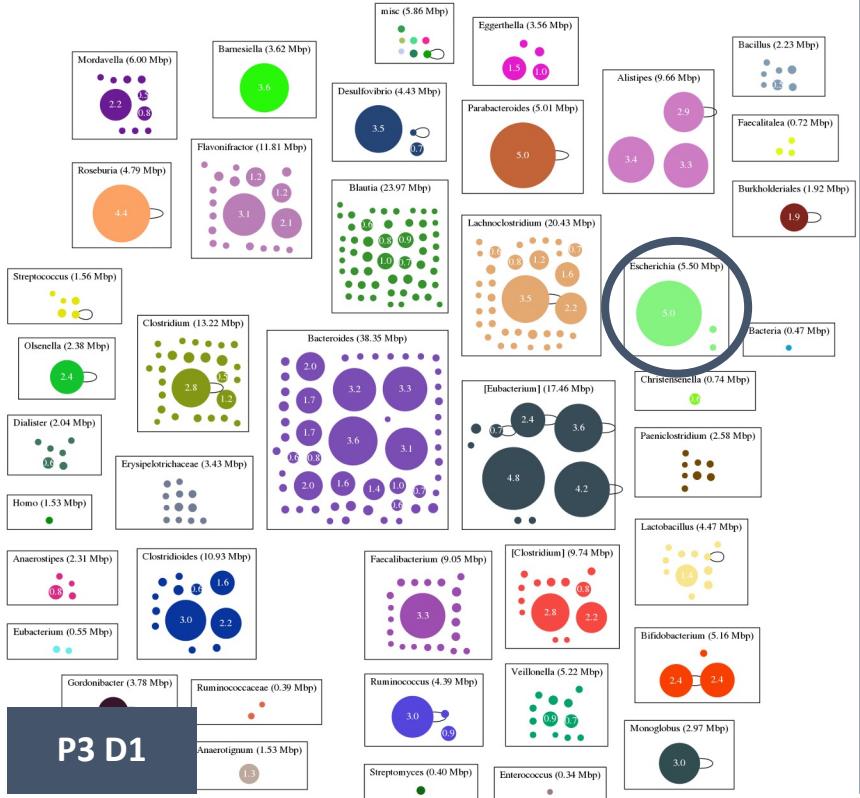
Reticulatus is our snakemake pipeline for routine long read metagenomics

steps are implemented as a chain of rules, snakemake infers the necessary order

reticulatus builds on top of this to provide an easy way to adjust parameters, pick software versions and set strategies for assembly and polishing

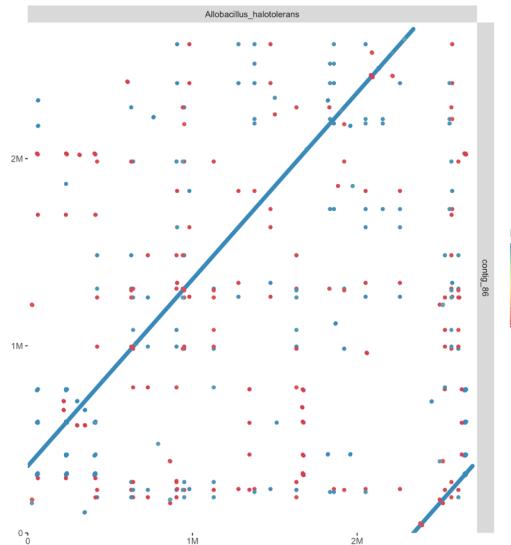


Spot the difference

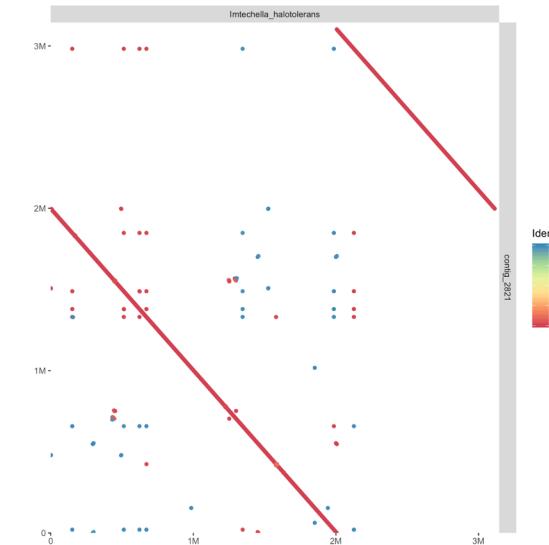


Long contigs, but are they any good?

Add spike-in control of unusual bacteria to faeces at <1% abundance to test

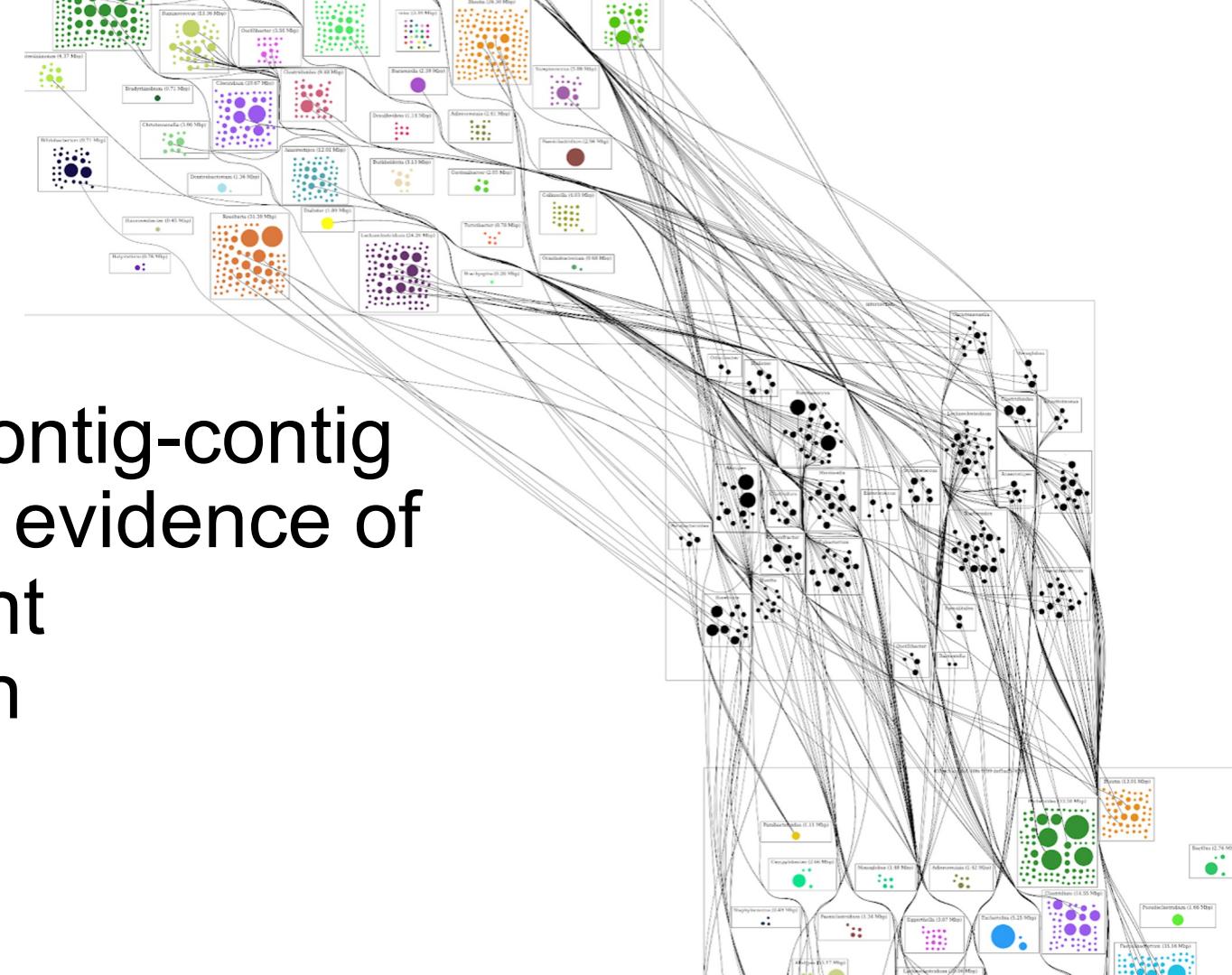


Allobacillus halotolerans



Imtechella halotolerans

Analysing contig-contig overlaps for evidence of donor-patient transmission



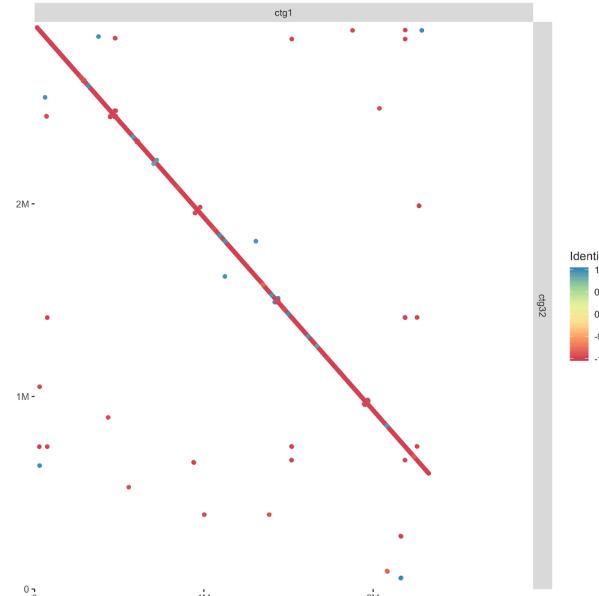
Inspecting candidate transmissions

Take donor stool and recipient after 56 days

Drop contigs <200kb

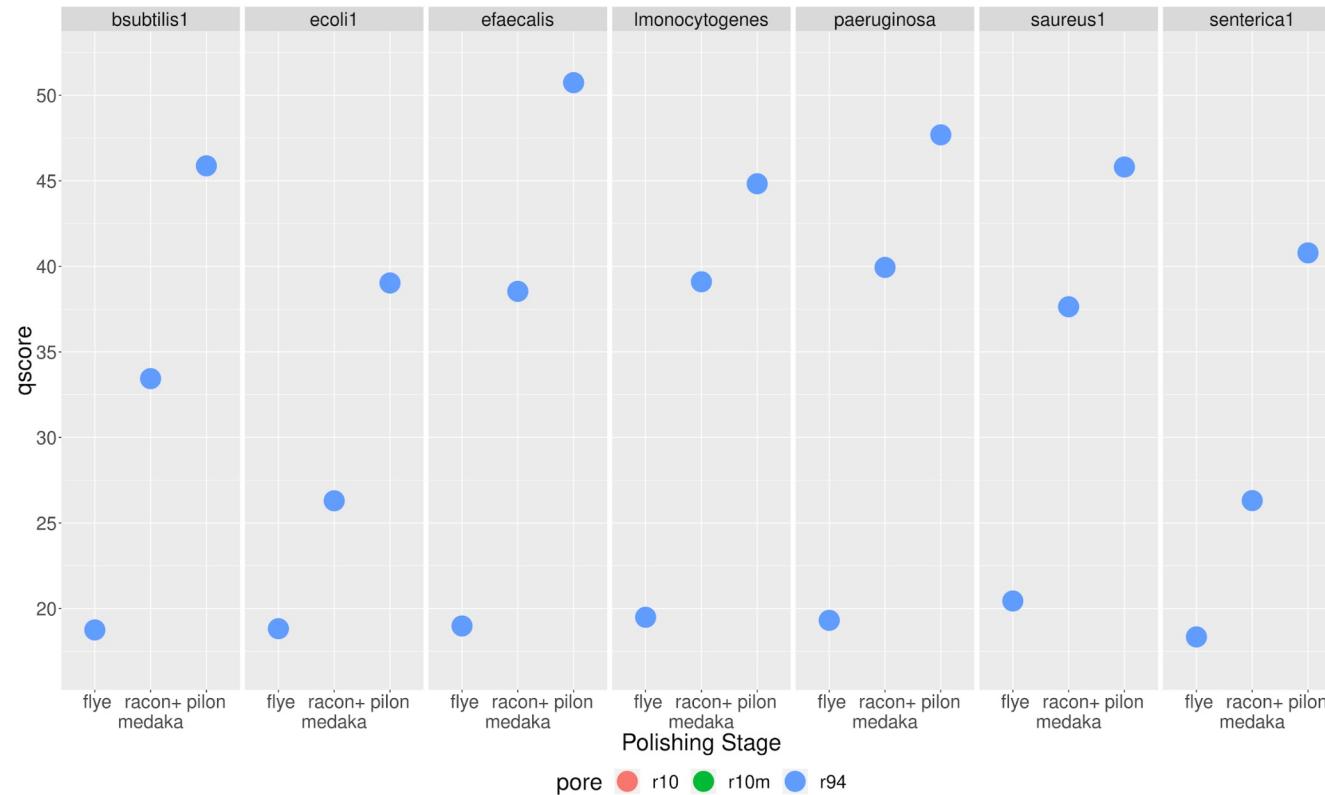
FastANI to look for high identity pairs (uses Mash)

Identifying completely novel taxa

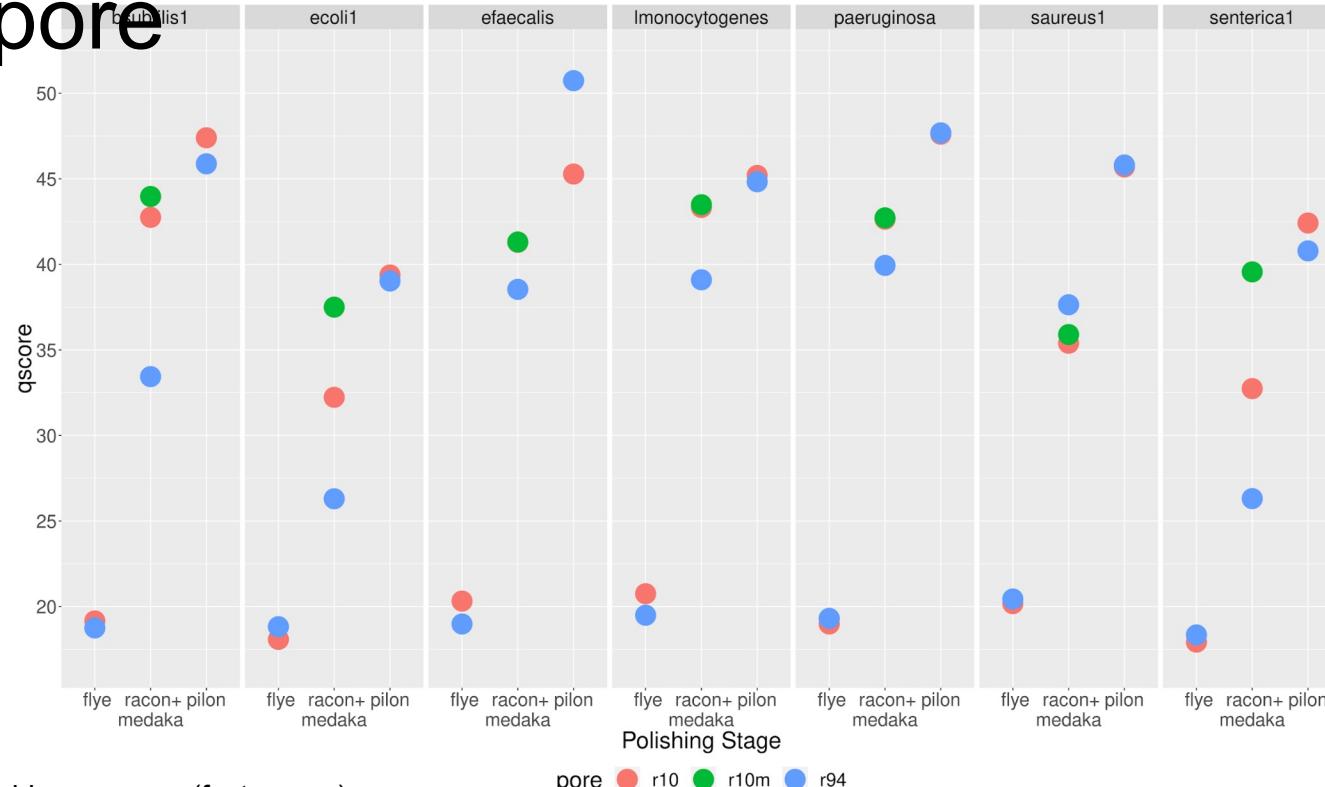


	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	Bacteroides xylosoxylans strain H207 chromosome, complete genome	427	579	99%	2e-114	64.47%	CP041230.1
<input checked="" type="checkbox"/>	Bacteroides fragilis strain DCMSKEJBY0001B chromosome, complete genome	427	486	99%	2e-114	64.47%	CP036546.1
<input checked="" type="checkbox"/>	Bacteroides thetaiotaomicron strain 7330, complete genome	427	537	99%	2e-114	64.47%	CP012937.1
<input checked="" type="checkbox"/>	Bacteroides thetaiotaomicron VPI-5482, complete genome	427	631	99%	2e-114	64.47%	AE015928.1
<input checked="" type="checkbox"/>	Bacteroides sp. A1C1 chromosome, complete genome	224	327	57%	4e-53	65.92%	CP036491.1

Long read polishing gets us most of the way



Higher consensus accuracy achievable with R10 pore



Conclusions

Long read metagenomics is a powerful tool for inspecting microbial communities

ONT platforms offer an opportunity to generate lots and lots of long reads

Read lengths are relatively short when using bead-beating and/or spin column based extractions

Metapolzyme was ineffective for lysing this strain of *Cryptococcus*

Optimised extraction methods can give very long reads and highly contiguous assemblies

Conclusion cont.

Yields on PromethION are sufficient for shotgun metagenomics for complex microbial communities (~500 ng input required)

Genome scale information recovered over a range of 4 logs difference in abundance, detection over 6 logs.

Evidence for transmissions from FMT donors to recipients including novel genomes

Long read assembly and polishing yields Q40-genomes

Acknowledgements

University of
Birmingham Nick Loman
 Josh Quick

CLIMB Radoslaw Poplawski

Zymo Research Shuiquan Tang

STOP-Colitis Tariq Iqbal
 Susan Manzoor
 Nabil Quraishi
 Peter Hawkey
 Birmingham Clinical Trials Unit
 Clinical Donors and Recipients

NVIDIA Mike Vella
 Fernanda Foertter

Oxford Nanopore Chris Wright

Bloomberg Daniel Evans

Thanks!

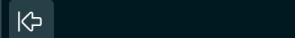
@samstudio8

[github.com / samstudio8](https://github.com/samstudio8) / reticulatus

**Ultra-deep, long-read nanopore sequencing of mock
microbial community standards**

Samuel M Nicholls, Joshua C Quick,
Shuiquan Tang, Nicholas J Loman
GigaScience, May 2019

Sequencing live demo



Dashboard

Getting started

Posts

Groups >

Nanopore Learning

Documentation

Highlights >

Updates

Support

LC Blog

NCM Blog

Notifications >

Profile >

DOCUMENTATION

Nanopore Documentation

DOCS

LSK114

[Home](#)[Plan](#)[Prepare](#)[Sequence](#)[Analyse](#)[Application Workflows](#)

[Ligation sequencing V14 - Direct cDNA sequencing \(SQK-LSK114\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 05/15/2023

[Telomere-to-telomere sequencing \(T2T\) on PromethION \(SQK-LSK114-XL and SQK-ULK114\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 06/07/2023

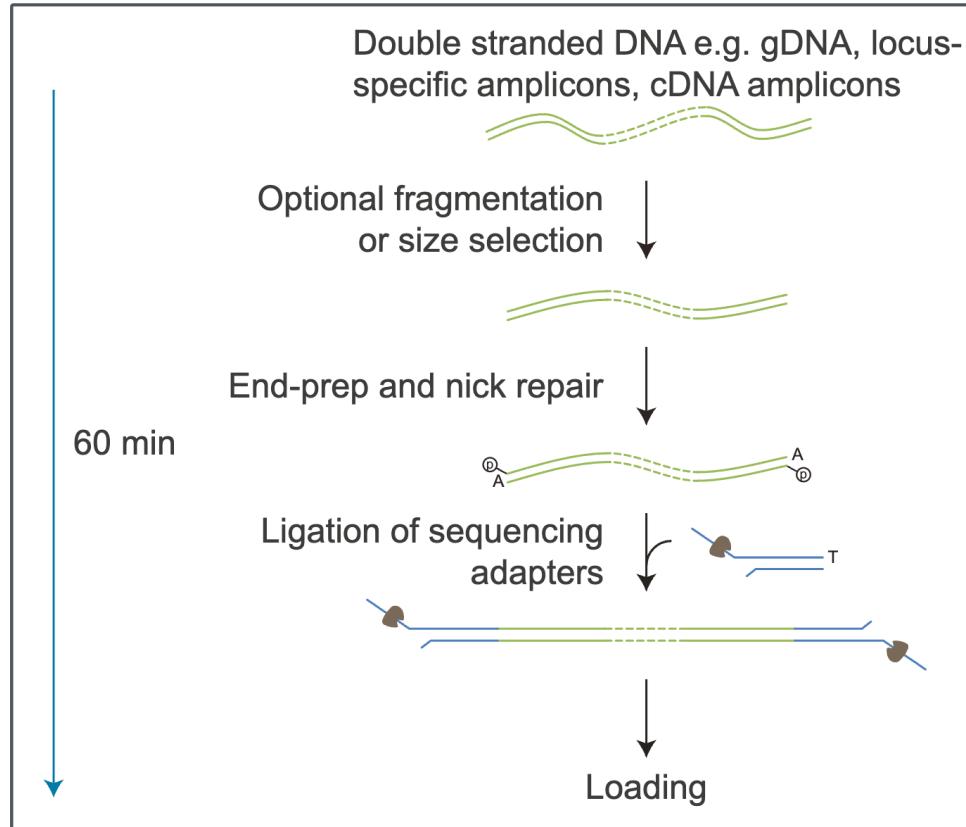
[Ligation sequencing V14 - low input by PCR \(SQK-LSK114 with EXP-PCA001\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 05/15/2023

LSK114 overview

- 1 µg high molecular weight genomic DNA / 100-200 fmol amplicon DNA
- 100+ ng high molecular weight genomic DNA if performing DNA fragmentation



DNA sample

Table 1: Microbial Composition

Species	Theoretical Composition (%)				
	Genomic DNA	16S Only ¹	16S & 18S ¹	Genome Copy ²	Cell Number ³
<i>Pseudomonas aeruginosa</i>	14	5.1	4.6	7.8	7.9
<i>Escherichia coli</i>	14	12.4	11.2	10.9	11.0
<i>Salmonella enterica</i>	14	12.7	11.4	11.2	11.2
<i>Enterococcus faecalis</i>	14	12.1	10.9	18.8	18.8
<i>Staphylococcus aureus</i>	14	19	17.1	19.6	19.6
<i>Listeria monocytogenes</i>	14	17.3	15.6	17.8	17.9
<i>Bacillus subtilis</i>	14	21.4	19.2	13.2	13.2
<i>Saccharomyces cerevisiae</i>	2	NA	10	0.63	0.32

https://files.zymoresearch.com/protocols/_d6322_zymobiomics_hmw_dna_standard.pdf

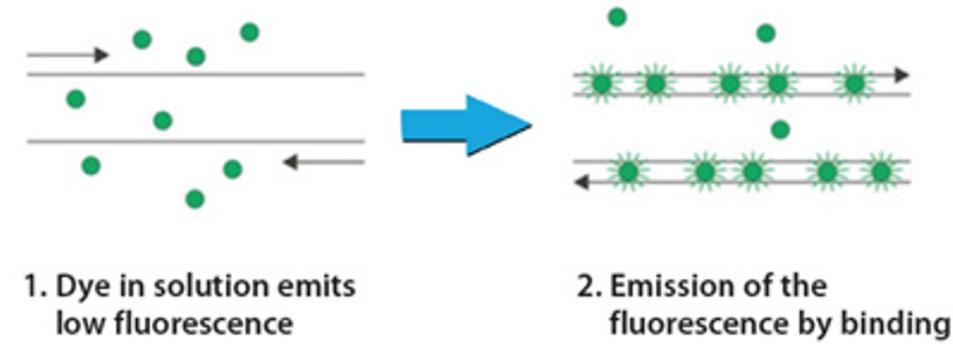
Product Specifications

- **Source:** seven bacteria (3 gram-negative and 4 gram-positive) and 1 yeast.
- **Reference genomes and 16S&18S rRNA genes:**
<https://s3.amazonaws.com/zymo-files/BioPool/D6322.refseq.zip>.
- **Storage solution:** 10 mM Tris-HCl and 0.1 mM EDTA, pH 8.0.
- **DNA concentration:** 100 ng/ μ l.
- **Impurity level:** < 0.01% foreign microbial DNA.
- **Average relative-abundance deviation:** <15%.
- **Microbial composition:** Table 1 shows the theoretical microbial composition of the standard.

The microbial composition of each lot was measured by shotgun metagenomic sequencing post mixing. The results (including the composition, impurities and abundance deviation) can be accessed through the Certificate of Analysis based on the lot number (printed on tube level) by the following link:
<http://www.zymoresearch.com/microbiomics/microbial-standards/zymobiomics-microbial-community-standards>.

Concentration

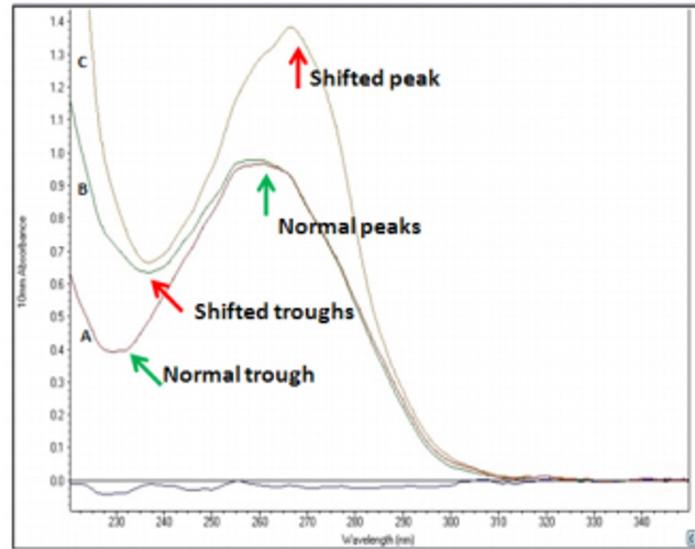
- Using nucleic acid stains e.g. SYBR Green I
- Highly specific to double stranded DNA
- Use standard curve to determine DNA concentration
- Qubit/fluorometer



<http://www.sigmaaldrich.com/technical-documents/protocols/biology/sybr-green-qpcr.html>

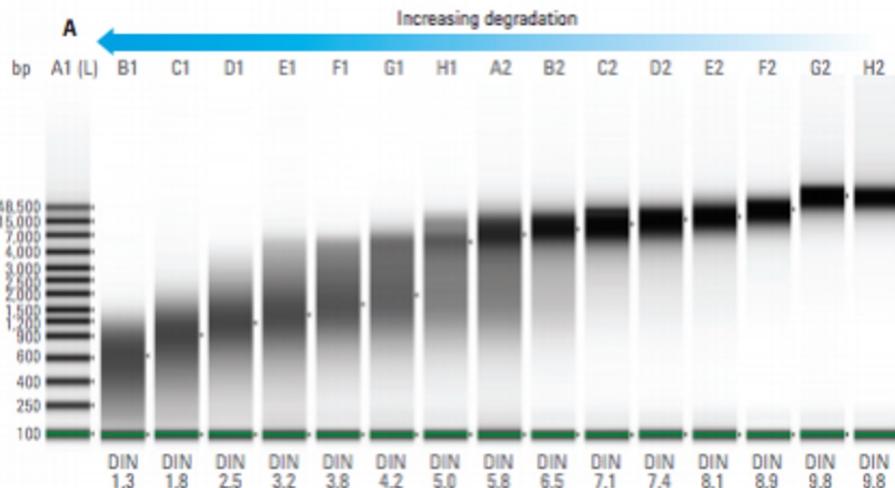
Absorbance

- Nucleic acids absorb at 260-270 nm
- Impurities may absorb at 230 and 280
- Important metrics:
 - 260/280 ideally 1.8
 - 260/230 ideally 2.0-2.2
- Salts, ethanol, phenol, protein



Fragment sizes

- Gel electrophoresis
- TapeStation DIN >9
- FemtoPulse
- Degradation caused by nucleases, freeze thaw, heating, pH, UV radiation



<http://www.agilent.com/cs/library/applications/5991-5258EN.pdf>

HMW standard TapeScreen

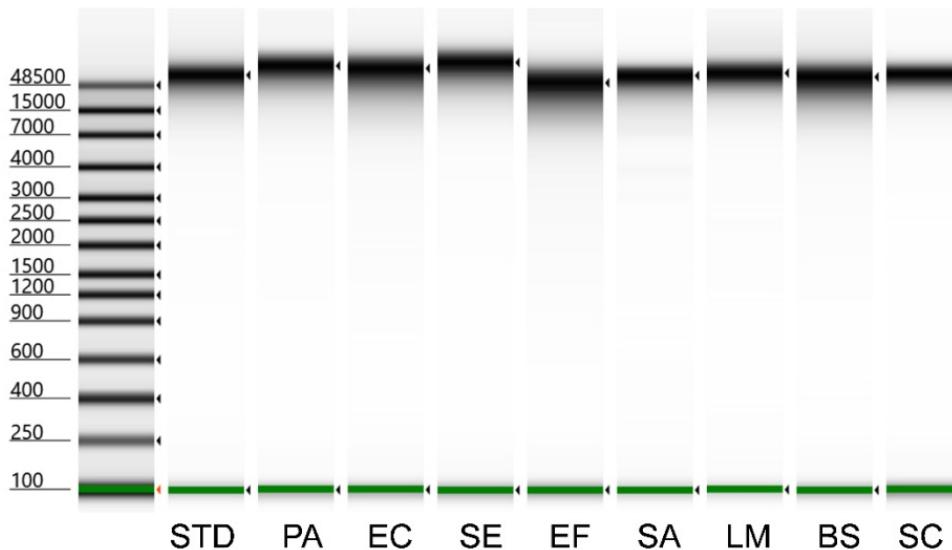


Figure 1. Genomic DNA from each species is >50 kb size. Sizing assessed by Agilent 2200 TapeStation® (Genomic DNA ScreenTape). The ZymoBIOMICS™ HMW DNA Standard (STD) is composed of the following organisms: (PA) *Pseudomonas aeruginosa*, (EC) *Escherichia coli*, (SE) *Salmonella enterica*, (EF) *Enterococcus faecalis*, (SA) *Staphylococcus aureus*, (LM) *Listeria monocytogenes*, (BS) *Bacillus subtilis*, (SC) *Saccharomyces cerevisiae*.

Input calculations

- Load 10-20 fmol of library
- To convert between mass and moles we need to know fragment size by running a gel
- Calculate molarity (e.g. 1 ug of 8kb)
 - $8000 \text{ bp} * 650 \text{ Da} = 5,200,000 \text{ Da}$
 - $1\text{E}^{-6} \text{ g} / 5,200,000 \text{ Da} = 1.92\text{E}^{13} \text{ mol}$
 - 192 fmol
- Calculate mass (e.g. 20 fmol, 48 kb)
 - $48,000 \text{ bp} * 650 = 31,200,000 \text{ Da}$
 - $2\text{E}^{-14} * 31,200,000 = 6.24 \times 10^{-7} \text{ g}$
 - 624 ng

MW dsDNA molecule (Da) = # of base pairs \times 650

$$\frac{\text{Mass}}{\text{Moles} \times \text{MW}}$$

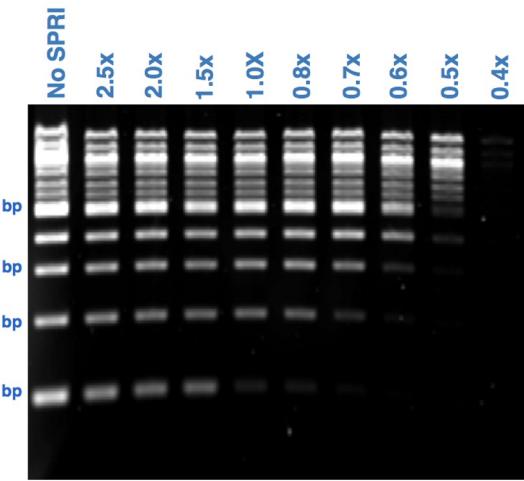
Shearing

- Optional but recommended for high yield
- Centrifugal shearing e.g. gTUBE (Covaris)
- Can also use 26-29G needle for larger fragments (20-50 kb)



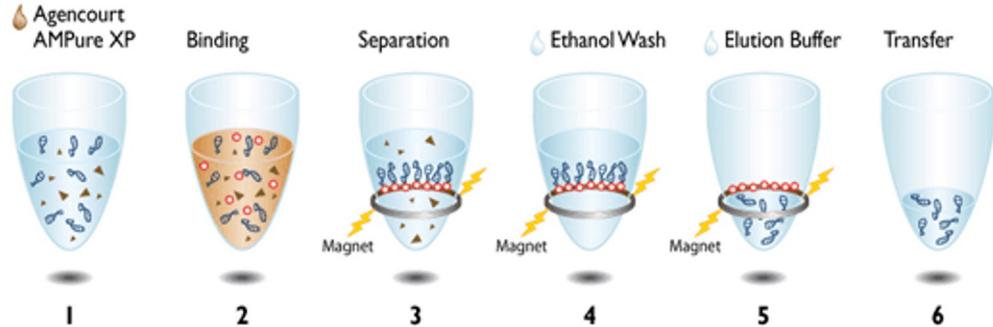
Size selection methods

- Ethanol typically used during washes >60% to keep bound
- Short fragments elute quicker than long fragments ‘Urban method’
- Using lower PEG/salt buffer can elute off DNA during the wash e.g. ‘LFB’



Ampure ratios using ladder

SPRI beads



- SPRI (Solid phase reversible immobilization) e.g. Ampure XP
- Condensed DNA will bind to beads
- The buffer contains PEG and NaCl to achieve the condensation
- Can be eluted by adding water
- Can perform size selection by changing the buffer concentration “ratios”

Prime flowcell

Reagent	Volume per flow cell
Flow Cell Flush (FCF)	1,170 µl
Bovine Serum Albumin (BSA) at 50 mg/ml	5 µl
Flow Cell Tether (FCT)	30 µl
Total volume	1,205 µl

https://community.nanoporetech.com/docs/prepare/library_prep_protocols/genomic-dna-by-ligation-sqk-lsk114/v/gde_9161_v114_revo_29jun2022/priming-and-loading-the-spot-on-flow-cell?devices=minion

Dilute Library

Reagent	Volume per flow cell
Flow Cell Flush (FCF)	1,170 µl
Bovine Serum Albumin (BSA) at 50 mg/ml	5 µl
Flow Cell Tether (FCT)	30 µl
Total volume	1,205 µl

https://community.nanoporetech.com/docs/prepare/library_prep_protocols/experiment-companion-minknow/v/mke_1013_v1_revcv_11apr2016/starting-a-sequencing-run-on-minion

tinyurl.com/nano-cinema23