Nanopore Sequencing in applied research at Agroscope

Andreas Bühlmann
Yvonne Bösch, Daniel Frei, Morgan Gueuning, Simon Blaser, Jürg Frey

Agroscope

Technology: Specifications vs. Real life

	Specifications	Agroscope
Instrument	MinION	MinION
Average read length	variable	~5kb (best = 19kb 2D)
Max read length	900kb	125kb (1D rapid)
Error Rate	5-15%	7% (1D)
Output	5Gb	2Gb (1D)
# reads	1M	1M
Instrument Price	1000$	1000$
Run Price	500- 900$	790$

https://blog.genohub.com/2017/06/16/pacbio-vs-oxford-nanopore-sequencing/

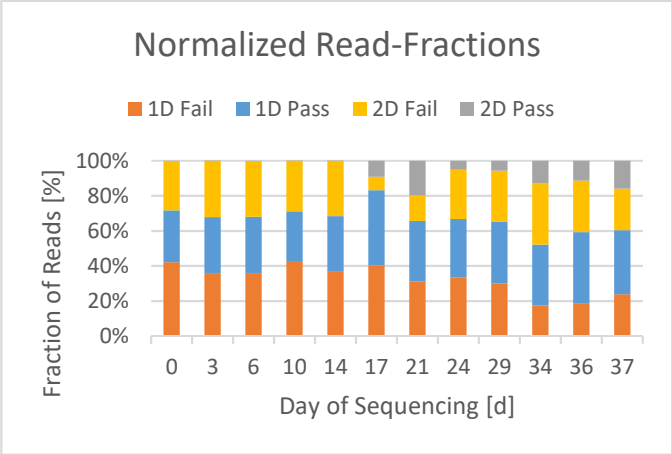
COI Identification of insects

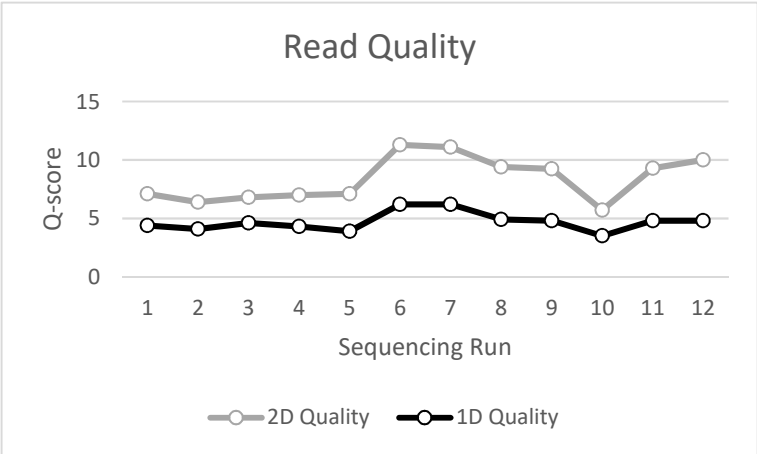
- The spread of harmful pest species by global trade can lead to important economic losses in agriculture
- Within the EU (ratified by CH), harmful plant pests are regulated as quarantine organisms (QO) commodities are tested at point of entry (PoE)
- Current situation: Sampling at Airport, Sanger Seq at Agroscope or Molecular Test directly performed at the Airport – Problem – new Species, Biotypes etc
- Goal: Replace sanger seq/molecular tests with Nanopore

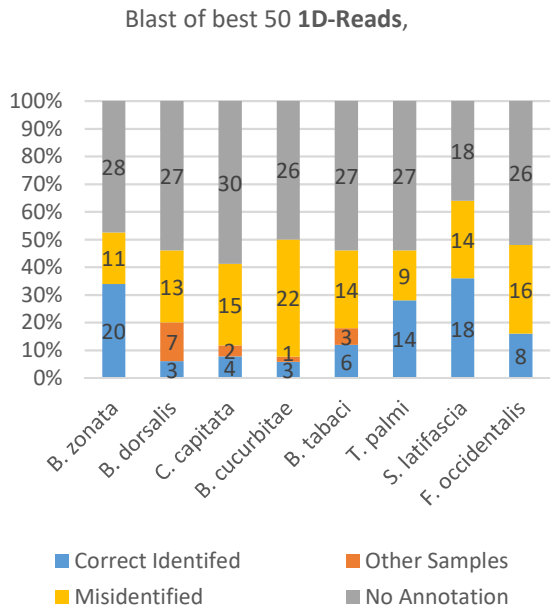


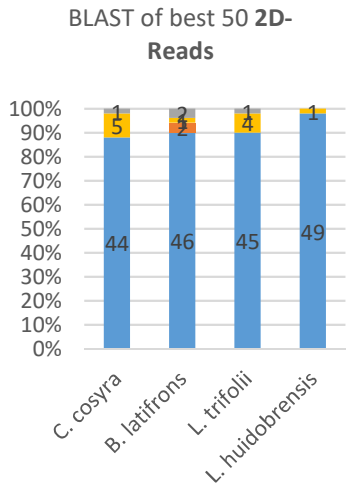

Source: www.agroscope.ch, kalamunda.wa.gov.au
H. Diem (SPPS), C. Heusser (Agroscope)

COI Identification of insects









Agroscope

COI Identification of insects

- Goal: Replace sanger seq with Nanopore
- Result: Possible, 12 libraries ok – try 12x 12 libraries?
- Result: Better than sanger
- Price a bit high – wait for SmidgION
- Work on dirty, low concentration DNA
- Regulatory stuff?

https://nanoporetech.com/products/smidgion

Source: www.agroscope.ch, kalamunda.wa.gov.au
H. Diem (SPPS), C. Heusser (Agroscope)

Plasmids harboring antibiotic resistance genes

- Collaboration with Maynooth University, IE
- Screen environmetal isolates for antibiotic resistance
- Clone plasmids into E.coli K12
- Isolate DNA, sequence on ONT

- Goal: Can we close plasmids and identify antibiotic resistance genes on ONT

Agroscope

Sequencing

Nanopore reads

Library	reads	Mbp	mean length	mean quality
NB01	32707	166.35	5086	15.7
NB02	33308	168.1	5047	15.7
NB03	149284	341.83	2290	15.6
NB04	153662	276.89	1800	15.2
NB05	28724	89.61	3120	15.4
NB06	103867	221.21	2129	15.5
NB07	151869	31.79	2046	16
NB08	127667	255.78	2003	15.5
Total	**781088**	**1551.56**	**2940.125**	**15.575**

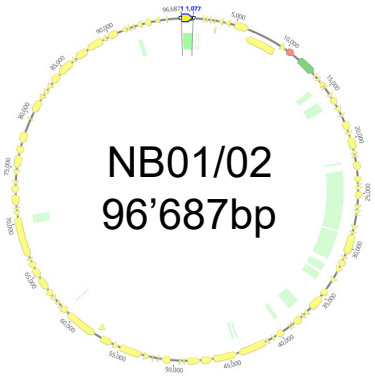
Illumina reads

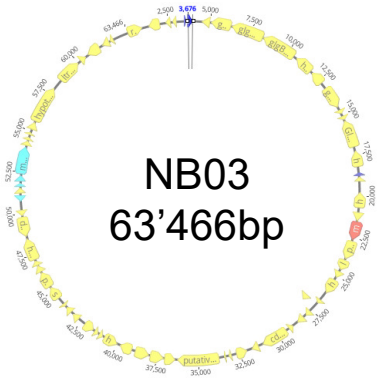
Library	reads	Mbp	mean length	mean quality
NB01	296766	36.54	123	36.3
NB02	395838	46.28	117	36.6
NB03	487244	55.93	115	36.4
NB04	190540	24.92	131	36.1
NB05	300610	32.34	108	36.6
NB06	312496	35.73	114	36.5
NB07	342648	39.19	114	36.6
NB08	257994	32.81	127	36.2
Total	**2584136**	**303.74**	**118.625**	**36.4125**

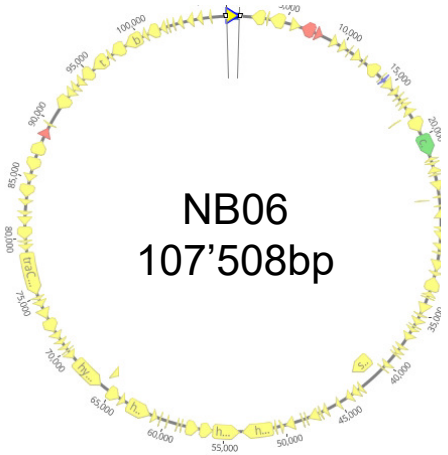
Assembly and annotation

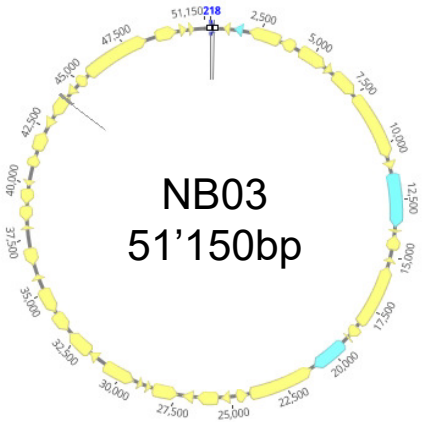
Assembly stats

library	contigs >5kb	after remapping	complete
NB01	12	7	1
NB02	5	5	1
NB03	85	6	2
NB04	7	5	3
NB05	4	3	2
NB06	312	1	1
NB07	64	64	0
NB08	119	1	1









Red: antibiotic resistance
Blue: toxin antitoxin
Cyan: heavy metal resistance
Green: colicin

Agroscope

Plasmids harboring antibiotic resistance genes

- Goal: Can we close plasmids and identify antibiotic resistance genes on ONT

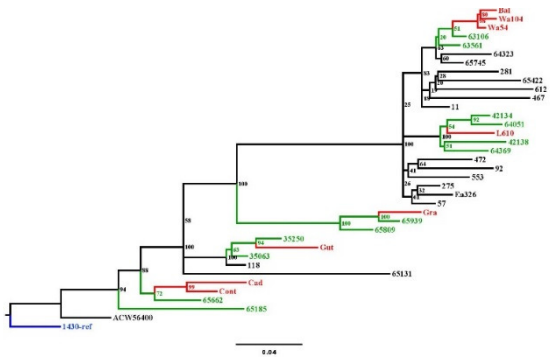
- Result: Close plasmids with ONT alone
- Advised to polish with Illumina at the moment
- Can identify resistance genes - novel?
- Try 1^2 as standalone

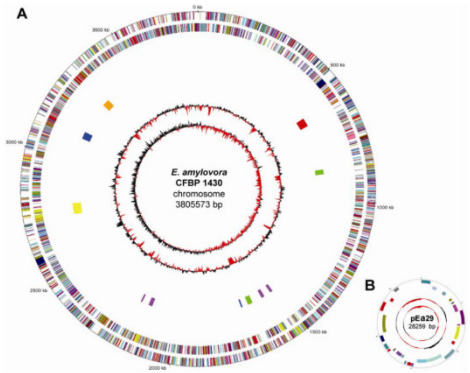
Agroscope

Bacterial genome sequencing

- *Erwinia amylovora*, causal agent of fire blight
- 3.9Mb genome, 1-3 Plasmids
- Good set of sequenced genomes in-house
- Need closed genomes for source tracking etc.

- Goal: can we close bacterial sequences with ONT

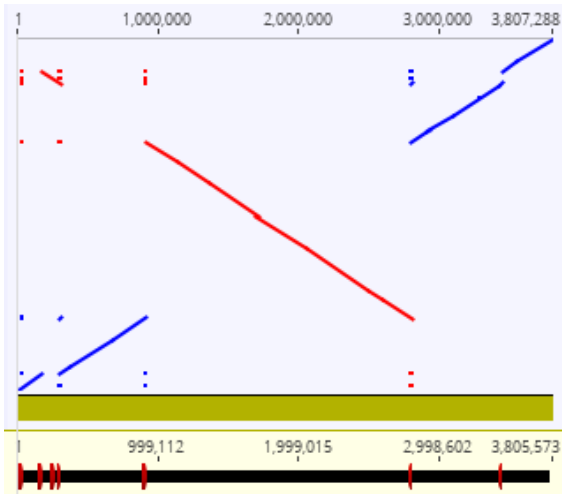


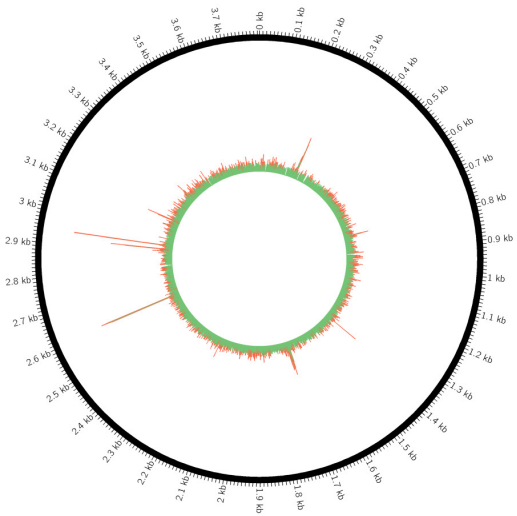


Sequencing and assembly

Library	Reads	Mbp	mean length	mean quality
EA UtrJ2	15185	67.85	3939	9.3
EA UtrJ2	4040	96.45	19200	13.1

	canu	unicycler
contigs	2	4
SNPs	2859	452
Indels and SNPs	30680	677
length	41334	1596
uncovered	2	8
total length	2	3402
Summe SNV+ uncovered	41336	4998
percent identity	98.91380352	99.8686663
known deletions		
Summe SNV+ uncovered	37520	1182
percent identity	99.01407751	99.9689403





Agroscope

Bacterial genome sequencing

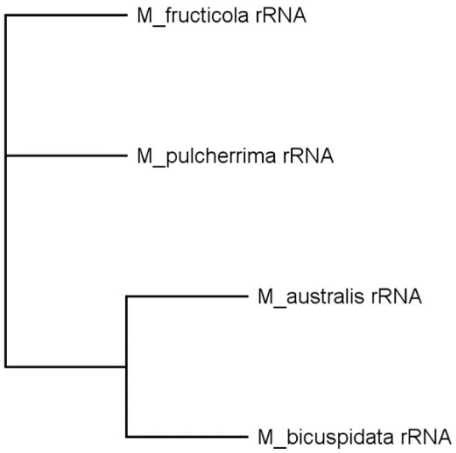
- Goal: can we close bacterial sequences with ONT

- Result: Close bacterial chromosomes with ONT alone
- Advised to polish with Illumina at the moment
- Unicycler works like a charm! Also with plasmids
- Try 1^2 as standalone?

Yeast genome sequencing

- Apple flower isolate of *Metschnikowia pulcherrima*
- *In-vitro* efficacy against plant pathogenic fungi
- Reference genome sequenced on PacBio in-house (Freimoser, Ahrens et al.)
- 14Mb genome, 15 chromosomes (PacBio)

- Goal: can we sequence a more complex genome



Yeast genome sequencing

- Extract DNA using Qiagen genomic Tips or Phenol Chloroform
- Library prep using SQK-LSK108, SQK-LSK208, and SQK-RAD002
- Sequencing on Nanopore chips R9.4 (so far 6 pieces)

- *De novo* assembly in Canu (Unicycler doesn't work here)
- Hybrid error correction in Pilon

	Canu assembly	Pilon correction
Contigs	13	13
N50	3.62Mbp	3.65Mbp
SNVs[1]	4800	2508
Insertions[1]	52597	1084
Deletions[1]	47814	494
%Identity[1]	99.071	99.899

[1]Comparison to inhouse PacBio assembly

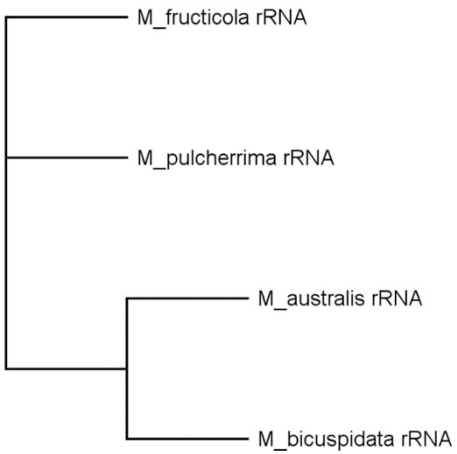
Species	Contigs	Genome size MB	G+C content %
Metschnikowia pulcherrima	13	16.1	45.7
Metschnikowia australis	154	14.4	47.2
Metschnikowia bicuspidata bicus.	48	16.1	47.9
Metschnikowia fruticola	93	26.1	45.9

Agroscope

Yeast genome sequencing

- Goal: can we sequence a more complex genome

- Result: We get a genome similar to the PacBio reference
- But, 2 contigs missing, 0.1% errors, still homopolymer errors
- Try one more chip…..



Metagenomics

Goal: Instead of sequencing 16S, which misses species and strain level, sequence whole rRNA

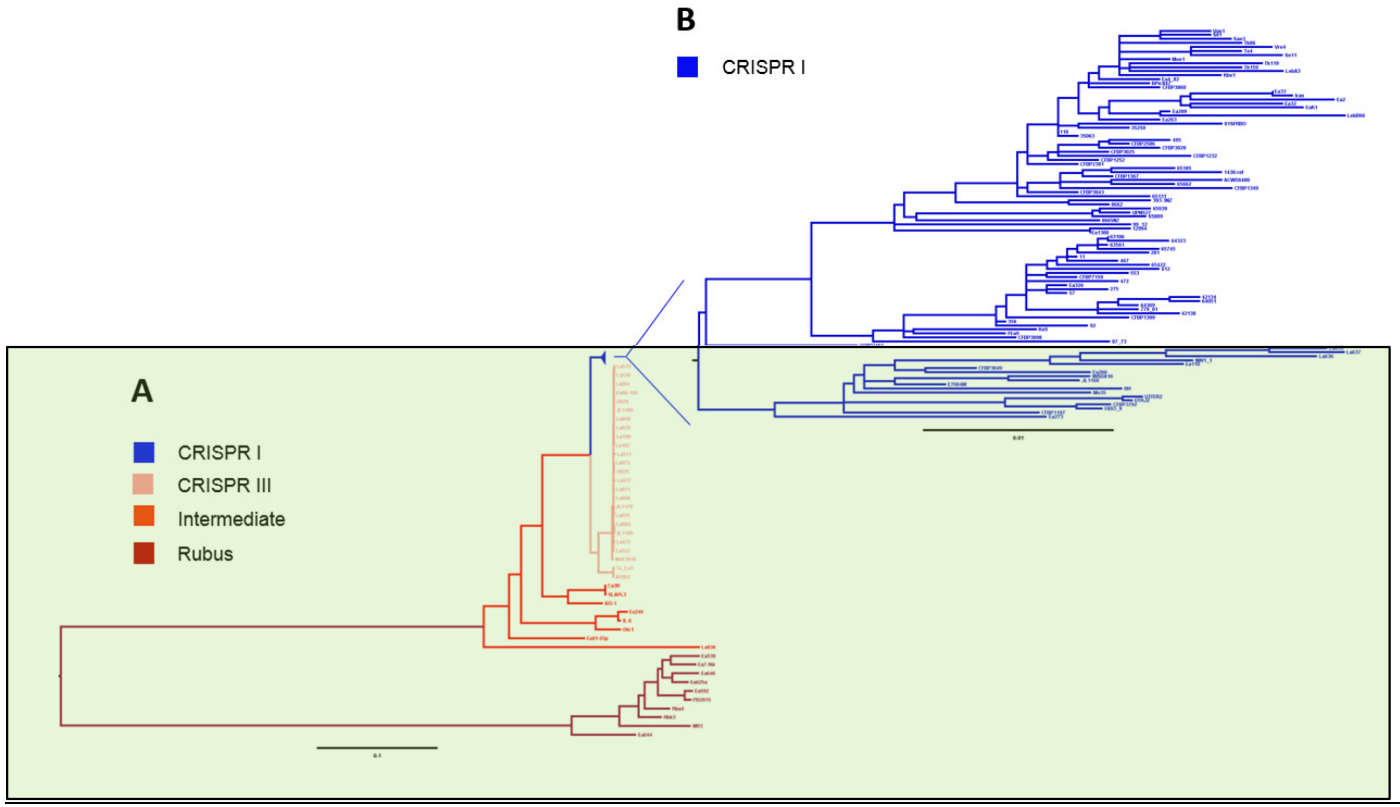
«State of the art»

«Shin-2016»

«Goal»

Metagenomics

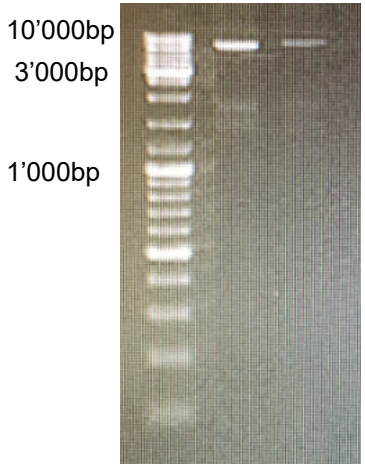
Example *E. amylovora*, should be able to distinguish between most strains

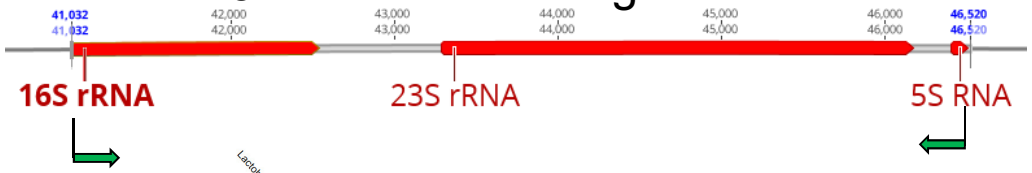


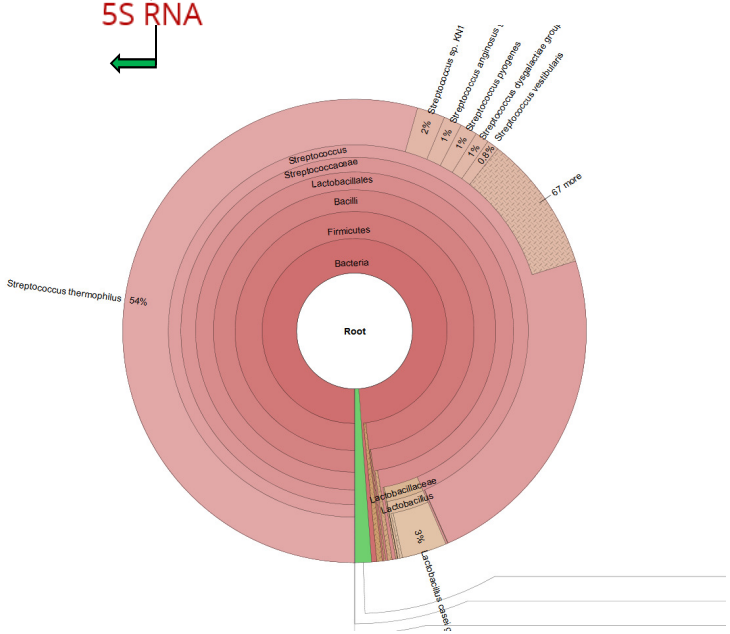
Agroscope

Metagenomics

- Mock mixes Lp, St, 10 strains each
- AMP-PCR Qiagen Long Range
- Expected Length ~5500bp
- 76% reads 4000-6000bp
- >85% hit correct genus
- ~5% barcode leakage





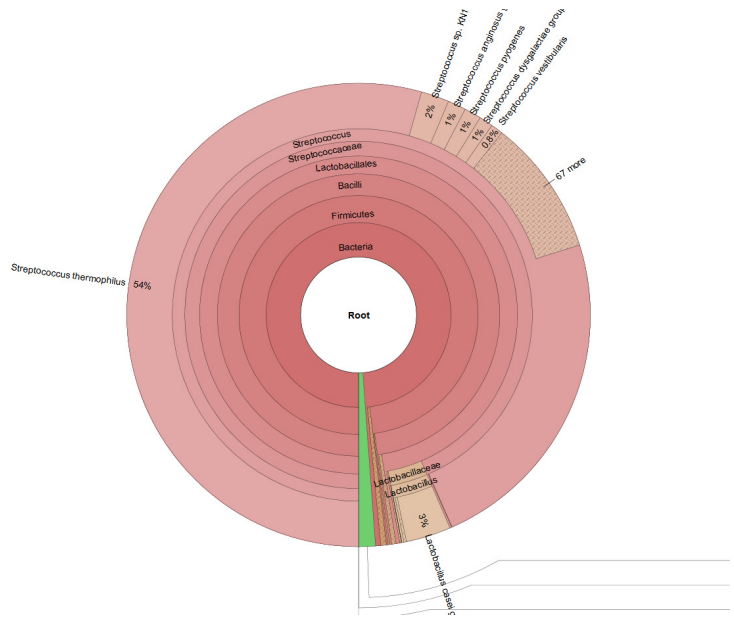


Metagenomics

Goal: Instead of sequencing 16S, which misses species and strain level, sequence whole rRNA

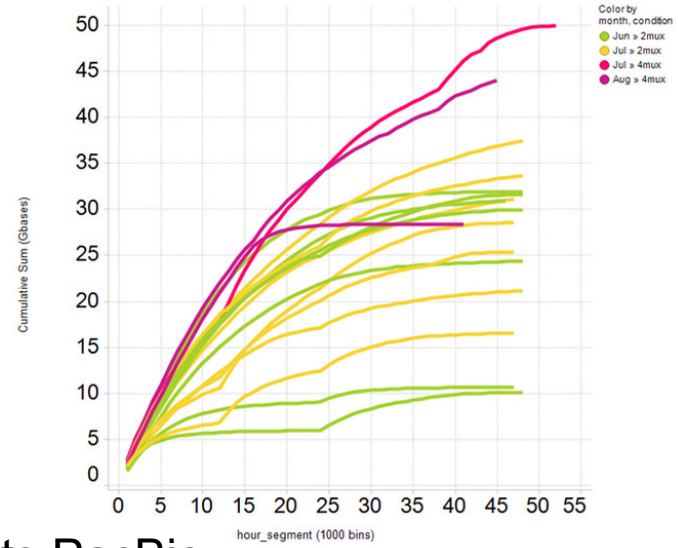
Result: Ongoing

- Reference sequences are absent
- Uclust clustering finds 10-15 clusters
- Reads map to clusters, how relevant?



Agroscope

Outlook and Future



- Direct RNA sequencing
- 5mC calling, 5mA?
- PromethIon
  - 50Gbp/chip
  - 2000$/chip – beats MiSeq, beats PacBio
  - 48 Chips/run
  - >2.5Tbp/run – beats HiSeq
- Bacterial population genomics, source tracking etc.
- Fungal, yeast comparative genomics
- Even go to plants….

Thanks

- Sonia Petignat and Team, Agroscope
- Fiona Walsh, University Maynooth
- Jürg Frey/ Christian Ahrens and Team, Agroscope/ SIB
- Florian Freimoser, Agroscope
- Marco Meola, Noam Shani, Agroscope
- Swiss Plant Protection, Service SPSS

Agroscope