Microbial outbreak surveillance system

Enabling automated microbial surveillance using Oxford Nanopore Technologies' long read sequencing
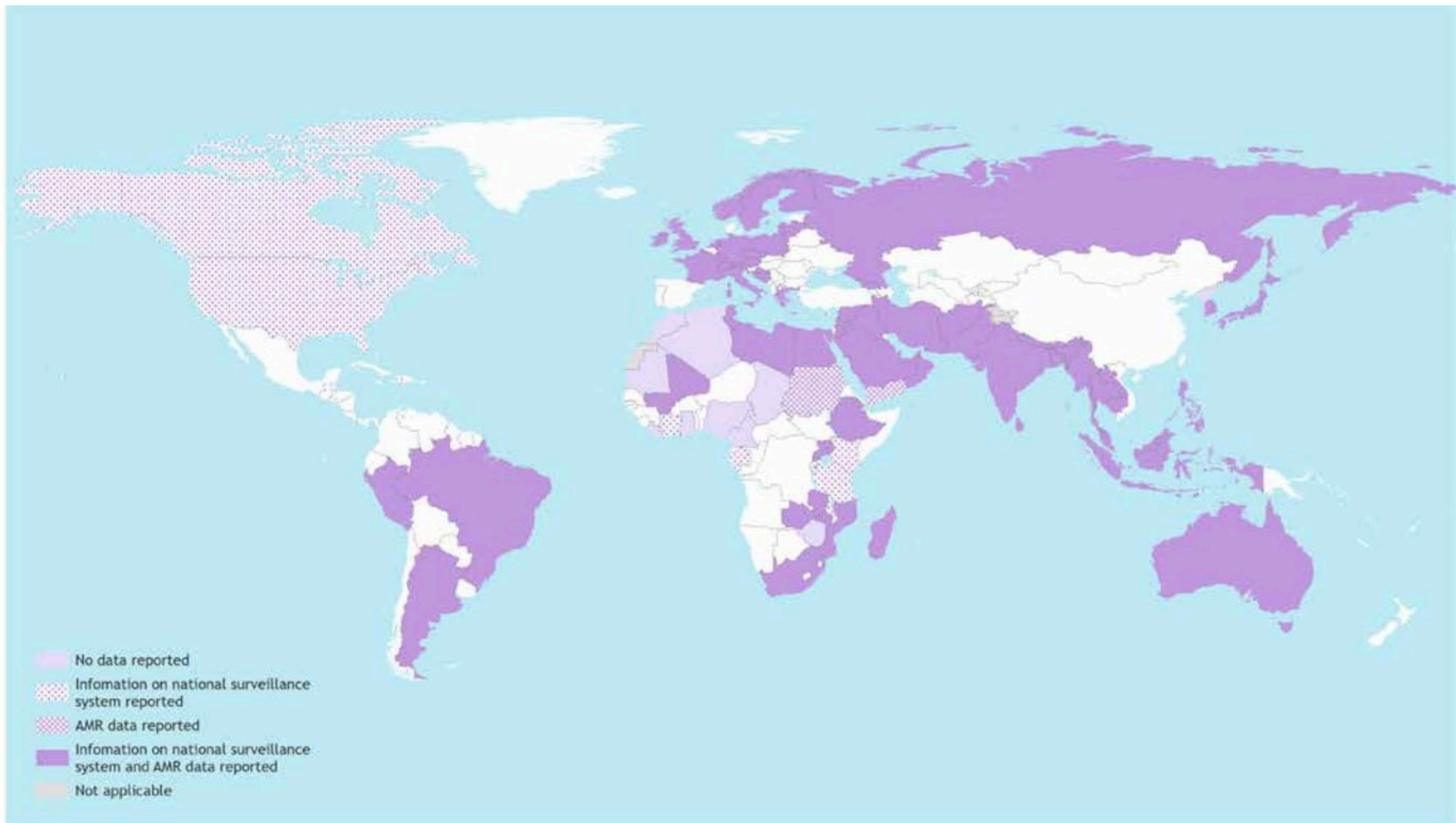
Outline for today

What the project is about

The science behind the project

What we have learnt so far

Current work and future ideas

	No data reported
	Infomation on national surveillance system reported
	AMR data reported
	Infomation on national surveillance system and AMR data reported
	Not applicable

Barriers to entry

- Huge labs
- Expensive equipment
- Large, trained staff
- Not automated report of results

The solution?

What needs to work scientifically

Phylogeny analysis (Hallgren et al. 2021)

Gene typing (Bortolaia V et al. 2020)

MLST typing

High scalability

What need to work practically?

- Localized sequencing
- Localized GPU base calling
- Localized computers
- Easy-to-use software for non-technical users
- No not just sequence everything and send it to the cloud?

The system

The back-end flow (Simplified)

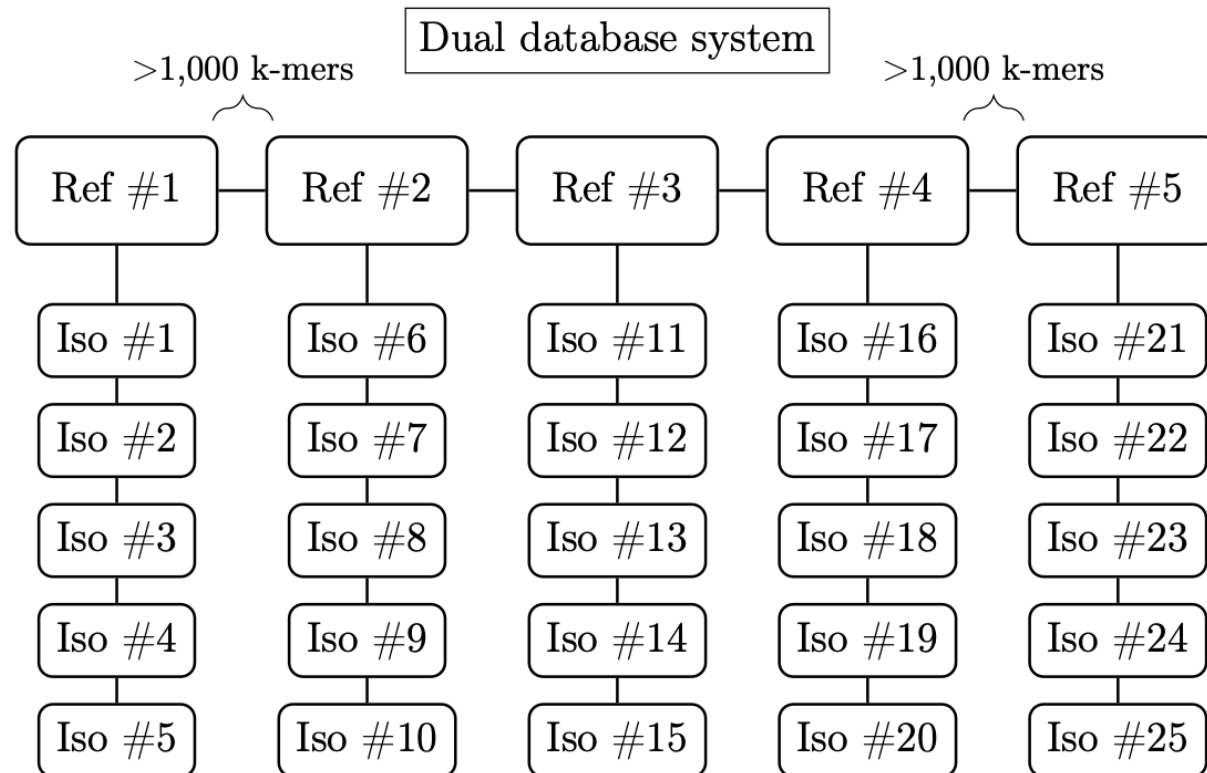|	
🫀	Type amr, virulence, plasmids, mlst
🔍	Search for reference
🔶	If reference found: kma alignment to reference MINTyper-based phylogenetic analysis
📄	Else: assemble to new reference
📄	Output PDF and report results

Database management

- Two databases:
  - A reference database

  - A metadata database

Table 3.2: Isolate samples from the CPO outbreak, which were aligned, and their identified references and travel destinations. $\delta* =$ SNP differences from the identified reference when using either MinION or Illumina reads.

E.coli isolate	$\delta*$ MinION	$\delta*$ Illumina	Reference identified	Patient travel destination
CPO20150014	71	59	NZ_CP024801.1	Thailand
CPO20150034	67	52	NZ_CP024801.1	India
CPO20150054	65	56	NZ_CP024801.1	Lebanon
CPO20160003	20	13	NZ_CP024801.1	Infected in Denmark
CPO20160077	20	14	NZ_CP024801.1	Infected in Denmark
CPO20180100	29	21	NZ_CP024801.1	Infected in Denmark
CPO20180105	29	21	NZ_CP024801.1	Infected in Denmark
CPO20180108	32	24	NZ_CP024801.1	Infected in Denmark
CPO20180119	34	26	NZ_CP024801.1	Infected in Denmark
CPO20180039	55	55	NZ_CP026473.1	Thailand
CPO20150011	146	95	NZ_CP029108.1	Pakistan
CPO20170014	98	93	NZ_CP029108.1	Pakistan

Challenge with deployment

Deployment must be extremely user friendly

Must technically be able to run on a laptop

GPU must be able to base call

= > Consumer targeted software engineering is hard, but we have learnt a lot!

It is impossible to predict pit falls if you can't imagine them!

Test driven development

A lot of technical debt to be cleared up

We are writing automated tests for every step of the pipeline

If system crashes, we will instantly know why/what happened.

Focus on speedy recovery rather than avoiding failures

Lesson learned: Once a laptop is shipped off to Africa, it becomes extremely difficult to identify and fix issues.

Future work

1	2	3	4
Staple release for microbial surveillance	Enable data sharing of results	Include virus workflow	Include metagenomics workflow