National Food Institute
Technical University of Denmark

Different suggestions for improving Global One Health surveillance of AMR

-

More data, better data, shared data

Frank M. Aarestrup

fmaa@food.dtu.dk, www.genomicepidemiology.org, www.globalsurveillance.eu, www.veo-europe.eu

What the world needs

- Real-time data on occurences of all infectious agents and AMR everywhere
  - Geography, reservoir and pathogen independent

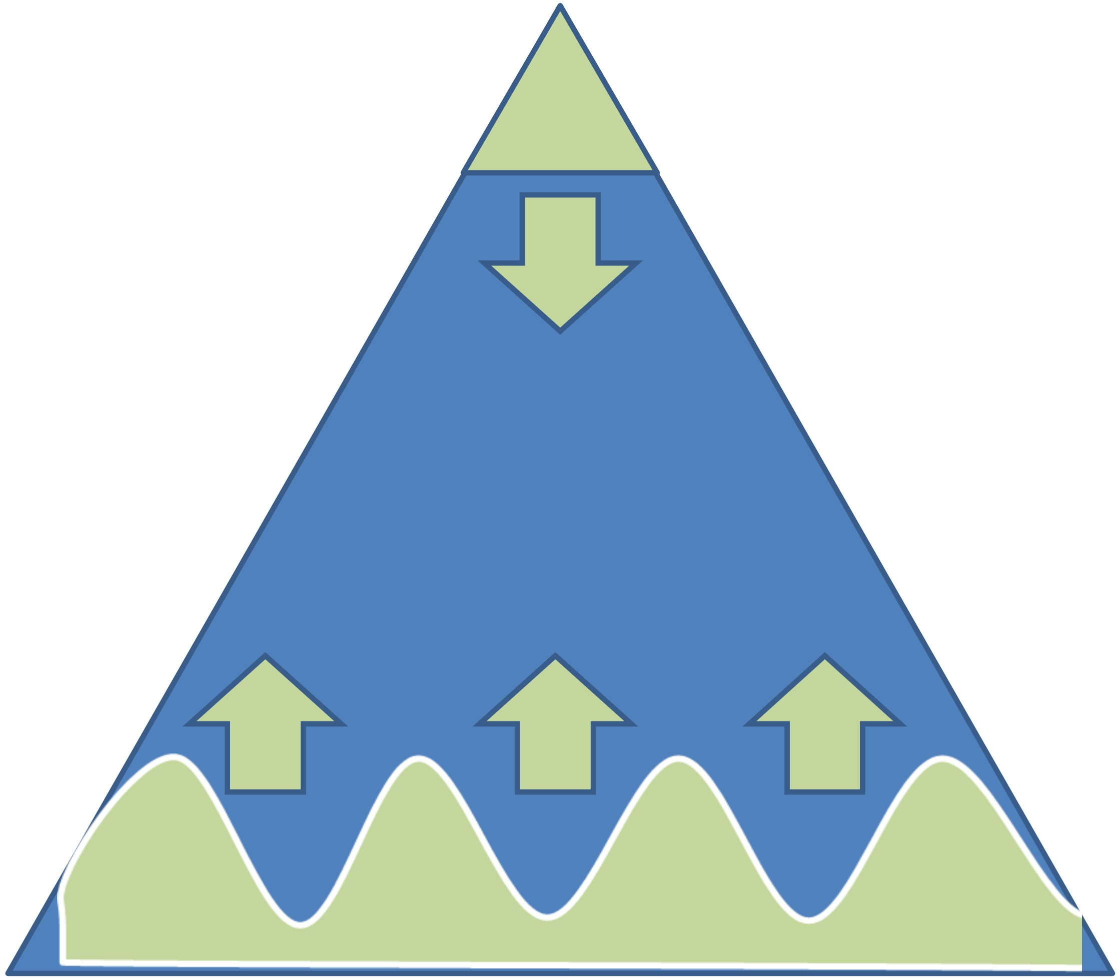
  - Observe trends and to rapidly compare between data
  - Transfer of information to those who need to:
    - Take public health respose
    - Develop tests and treatments
    - Take clinical decisions

IF WE DO NOT KNOW WHERE WE ARE WE HAVE NO IDEA WHERE WE ARE GOING

However, we still do not have a global surveillance

- What are the barriers?
  - Access to samples / isolates
  - Ability to generate data – phenotypic / genotypic
  - Ability to interpret data
  - Ability to share data – permissions / location

Top-down *or* bottom up?



Design and test the perfect system in a single or few countries and subsequently roll out across the globe

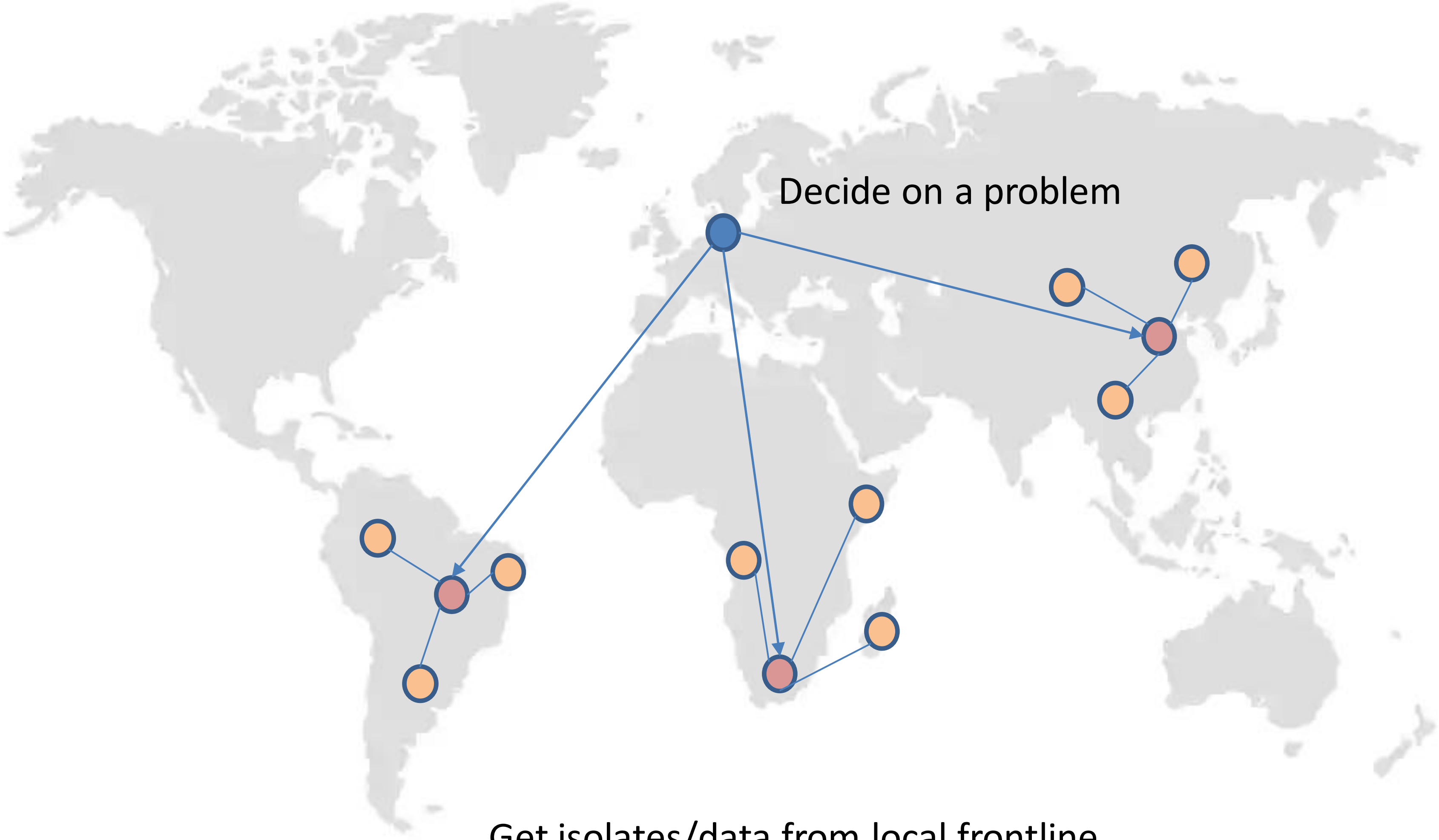
- Easy to document and publish
- Short-term

Empower countries / people where they are and hope it in general improves globally

- Difficult to document effects

Traditional – top-down

The WHO / ECDC / EFSA model



Decide on a problem

Get isolates/data from local frontline

Identify or establish reference centres

Build capacity for each individual test in each individual country

- Extremely expensive (2 million € per country, just for AMR)
- Limited flexibility
- Monopolizing at NRLs

Traditional – top-down

The Animal / Food / Human – not OneHealth model

WORLD ORGANISATION FOR ANIMAL HEALTH

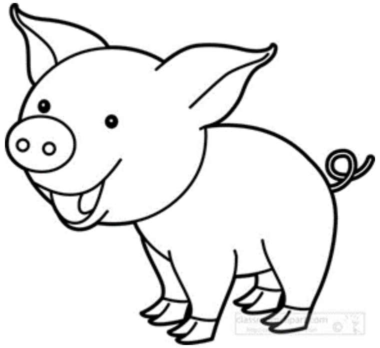
FIAT PANIS

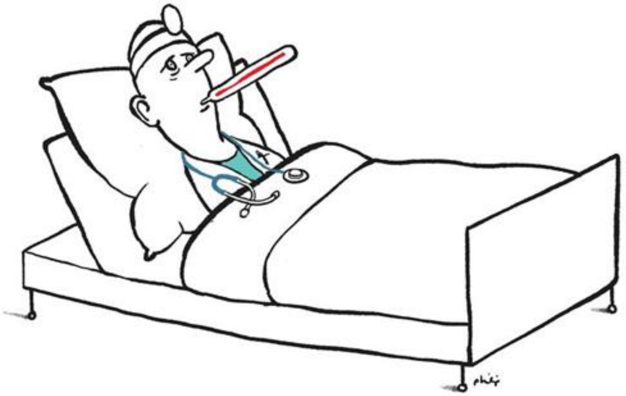
World Health Organization

USDA

CENTERS FOR DISEASE CONTROL AND PREVENTION

Some national attempts to create OneHealth institutions or collaboration





Identify or establish reference centres

Build capacity for each individual test in each individual country

- Extremely expensive
- Limited flexibility
- Monopolizing at NRLs

Has historically been difficult to build collaboration

Have – at times – been difficult to get data

Even more difficult on a global scale

Other options

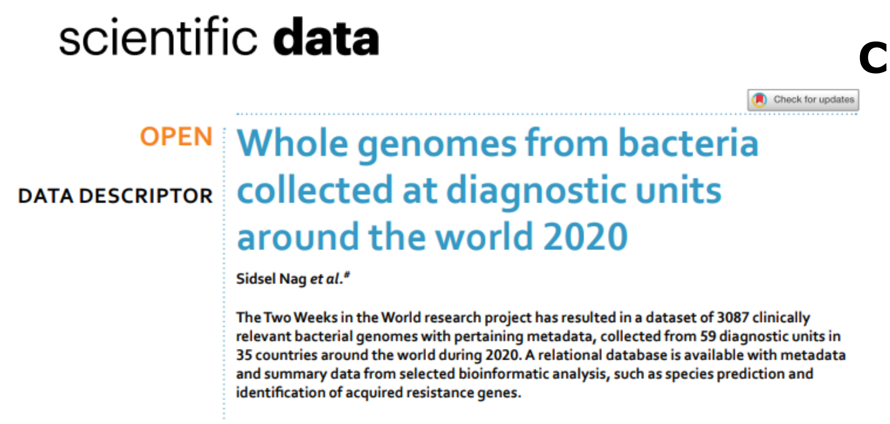
- **Perform**
  - Simply collect and analyse samples / isolates from the globe. Can be combined with training and capacity building

- **Enable**
  - Online tools
  - Frontline diagnostics and easy data sharing
  - Central private clouds (facebook of genomics)

- **Predict**
  - Improve data sharing options and utilize the randomly available data and build automatic models

Central collection of bacterial pathogens

A feasibility study



scientific **data**

Check for updates

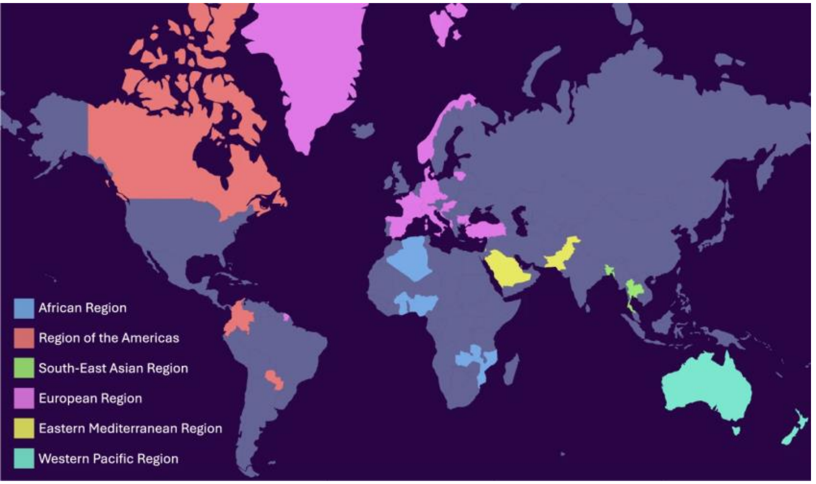
OPEN

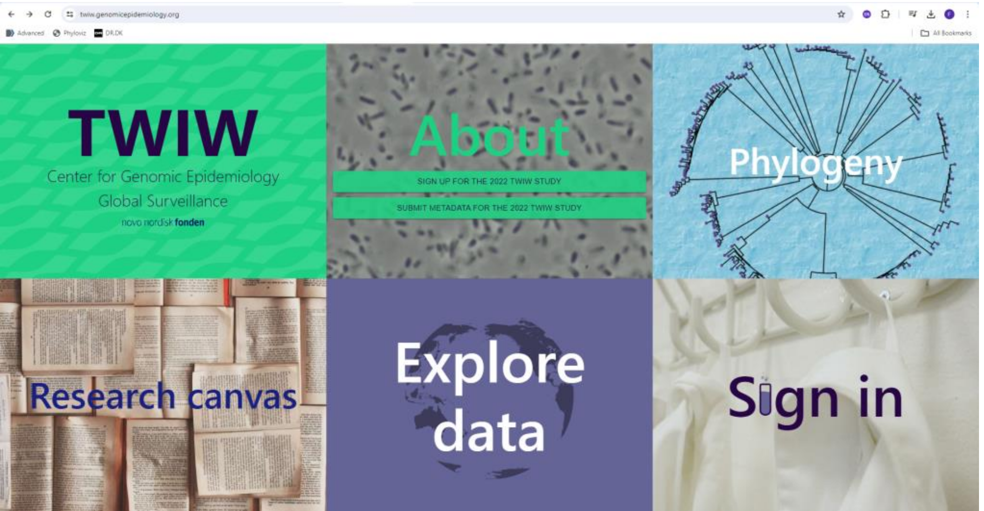
DATA DESCRIPTOR

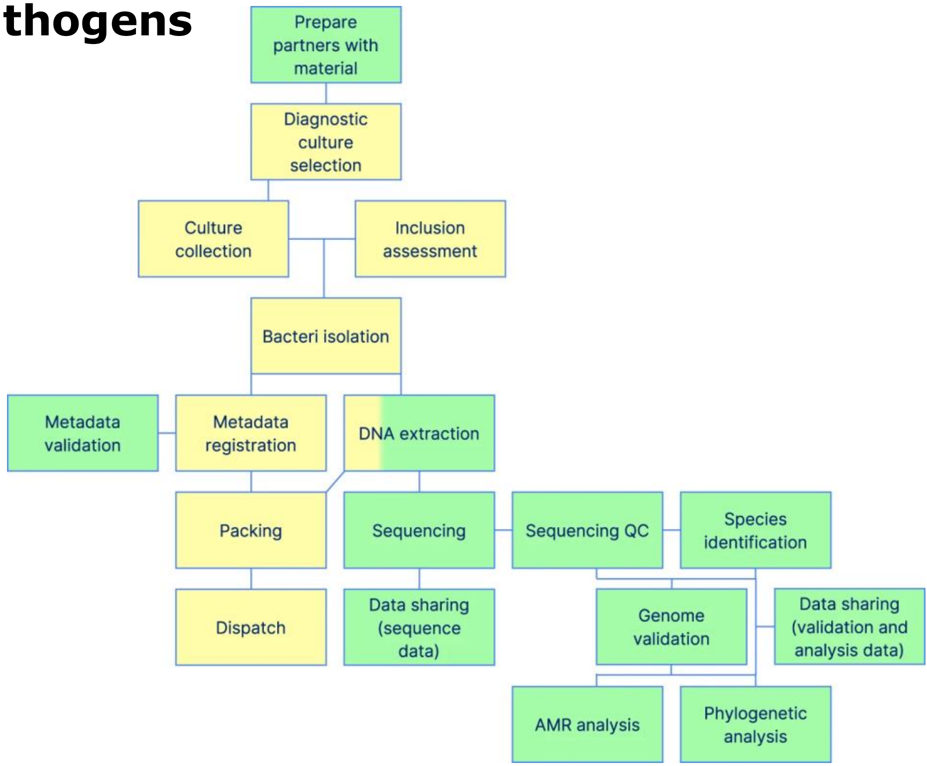
Whole genomes from bacteria collected at diagnostic units around the world 2020

Sidsel Nag *et al.*#

The Two Weeks in the World research project has resulted in a dataset of 3087 clinically relevant bacterial genomes with pertaining metadata, collected from 59 diagnostic units in 35 countries around the world during 2020. A relational database is available with metadata and summary data from selected bioinformatic analysis, such as species prediction and identification of acquired resistance genes.







Prepare partners with material

Diagnostic culture selection

Culture collection

Inclusion assessment

Bacteri isolation

Metadata validation

Metadata registration

DNA extraction

Packing

Sequencing

Sequencing QC

Species identification

Dispatch

Data sharing (sequence data)

Genome validation

Data sharing (validation and analysis data)

AMR analysis

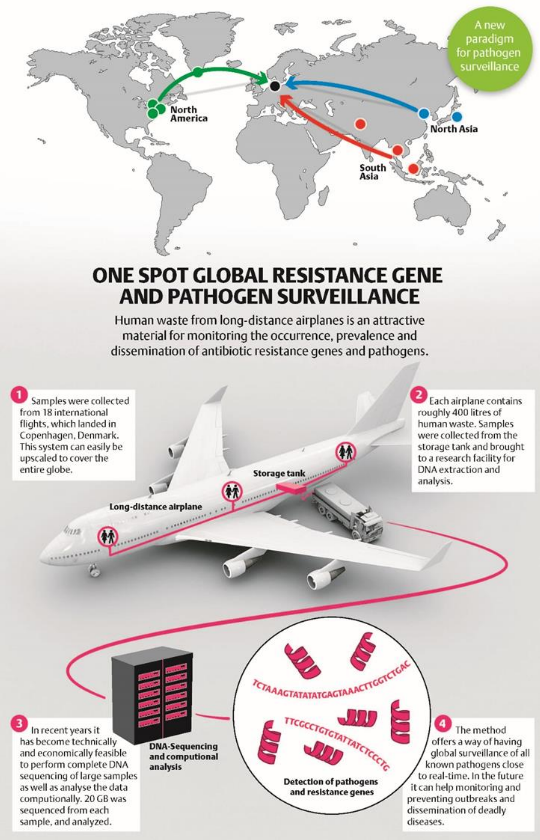
Phylogenetic analysis

- 59 diagnostic units from 35 countries
- Collected 3,087 “random” isolates
- Mix of shipping isolates or DNA -> Central sequencing
- Simple data openly available
- Raw data available for the participants
- A global solution would cost approx 1 million € per year

Metagenomics –
One technology that takes all

Metagenomics *is a methodology where we are purifying all DNA from all organisms in a sample and randomly sequencing small fragments. We can then both identify and quantify what was in that sample.*

Perform at global hot-spots



ONE SPOT GLOBAL RESISTANCE GENE AND PATHOGEN SURVEILLANCE

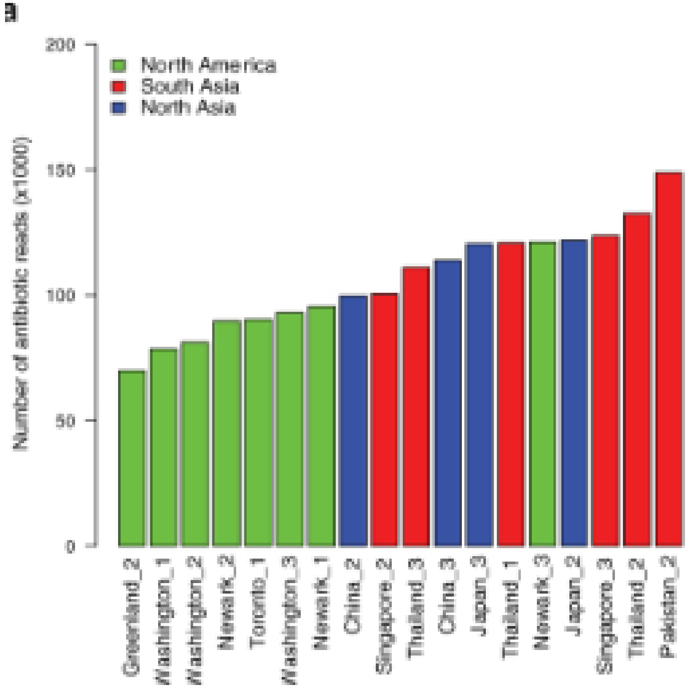
Human waste from long-distance airplanes is an attractive material for monitoring the occurrence, prevalence and dissemination of antibiotic resistance genes and pathogens.

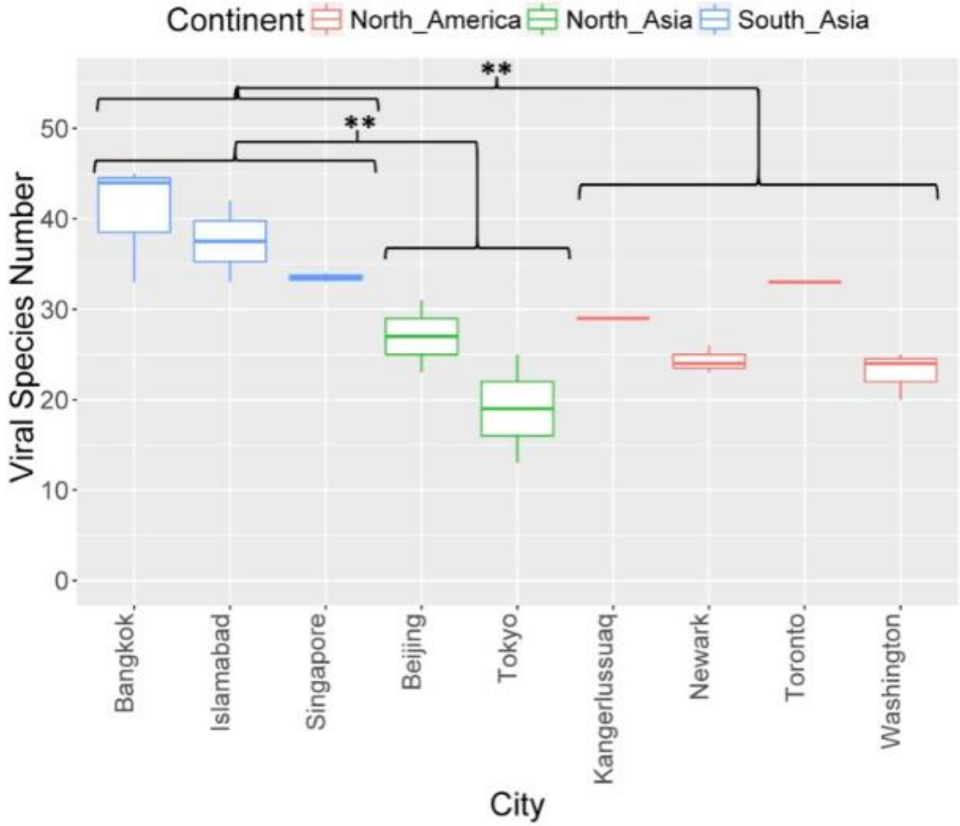
1. Samples were collected from 18 international flights, which landed in Copenhagen, Denmark. This system can easily be upscaled to cover the entire globe.

2. Each airplane contains roughly 400 litres of human waste. Samples were collected from the storage tank and brought to a research facility for DNA extraction and analysis.

3. In recent years it has become technically and economically feasible to perform complete DNA sequencing of large samples as well as analyse the data computionally. 20 GB was sequenced from each sample, and analyzed.

4. The method offers a way of having global surveillance of all known pathogens close to real-time. In the future it can help monitoring and preventing outbreaks and dissemination of deadly diseases.

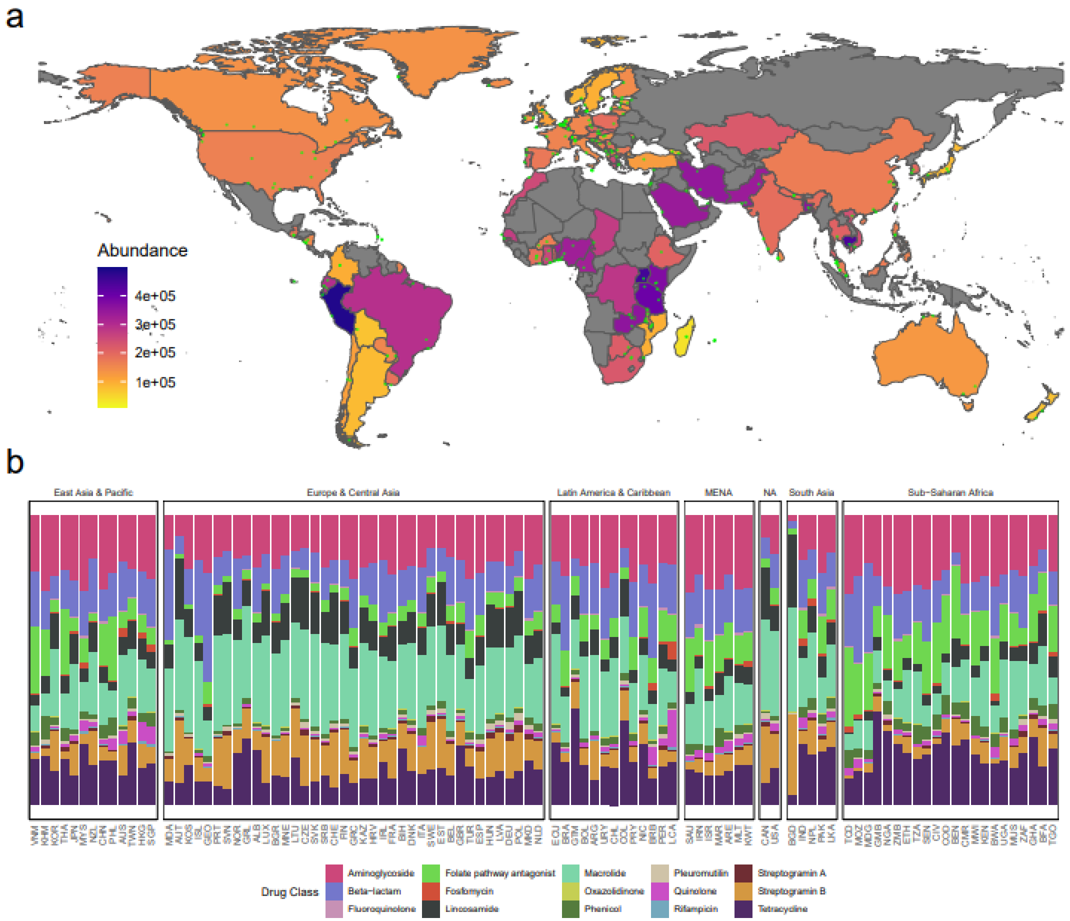


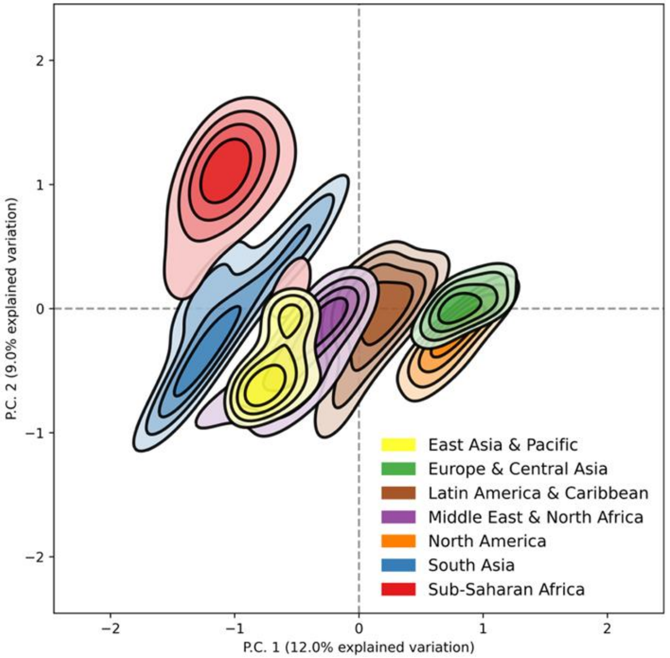


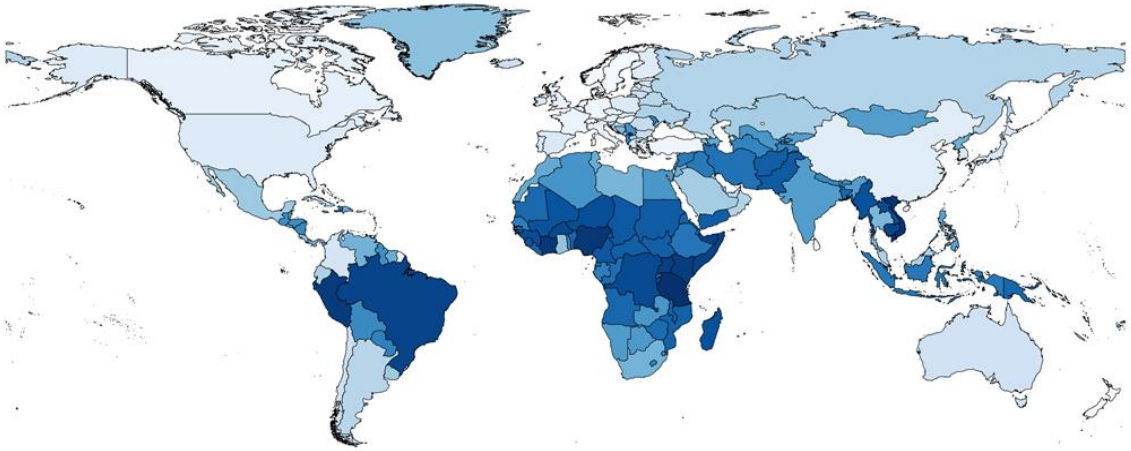
Pathogens and AMR

Pedersen et al. 2015, Hjelmsø et al. 2019

Surveillance using wastewater - AMR







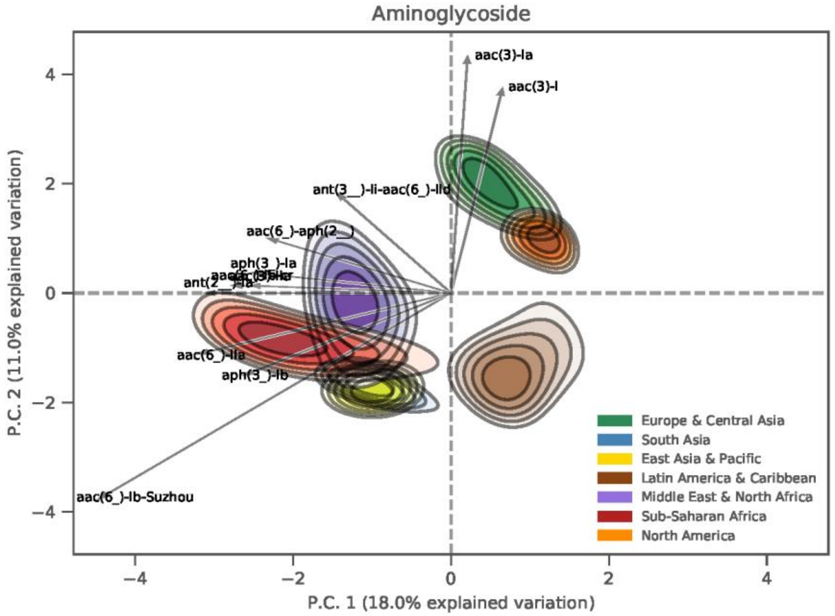
Reproducible

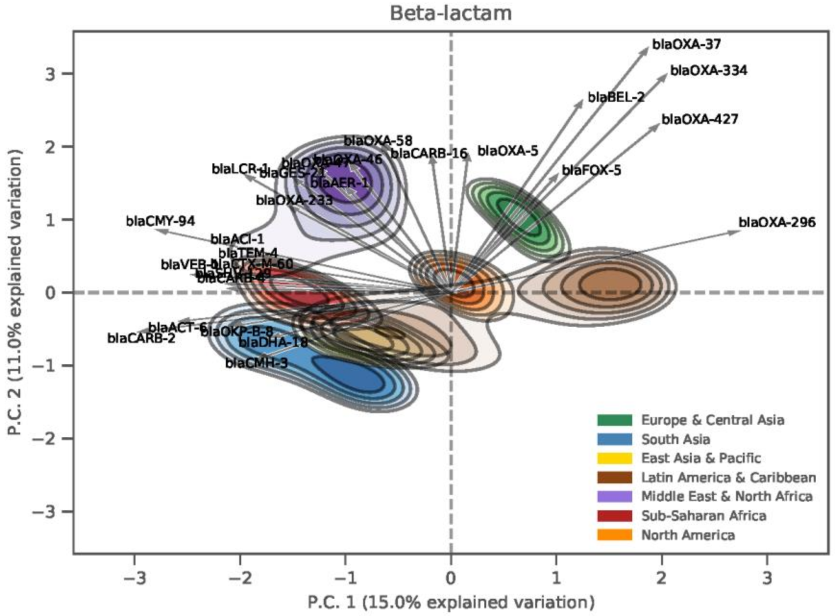
Clear separation of the world’s regions

Mainly driven by socio-economic factors

We can make predictions

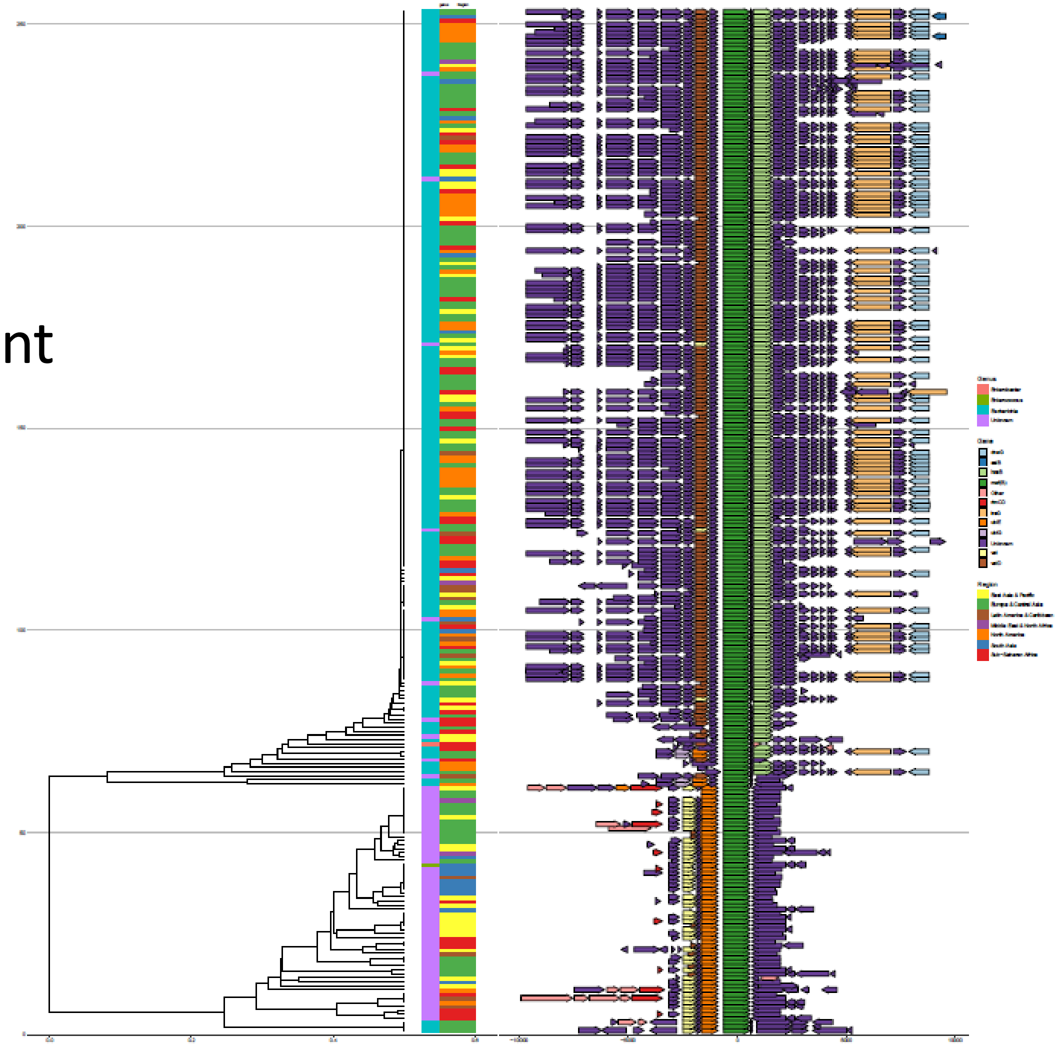
Hendriksen et al. Nature Comm. 2019
Aarestrup & Woolhouse. Science. 2020
Munk et al. Nature Comm. 2022





The different classes separate the world in different ways

Each ARG variant follow its own evolutionary trajectory



Comparing to socio-economic data from World bank

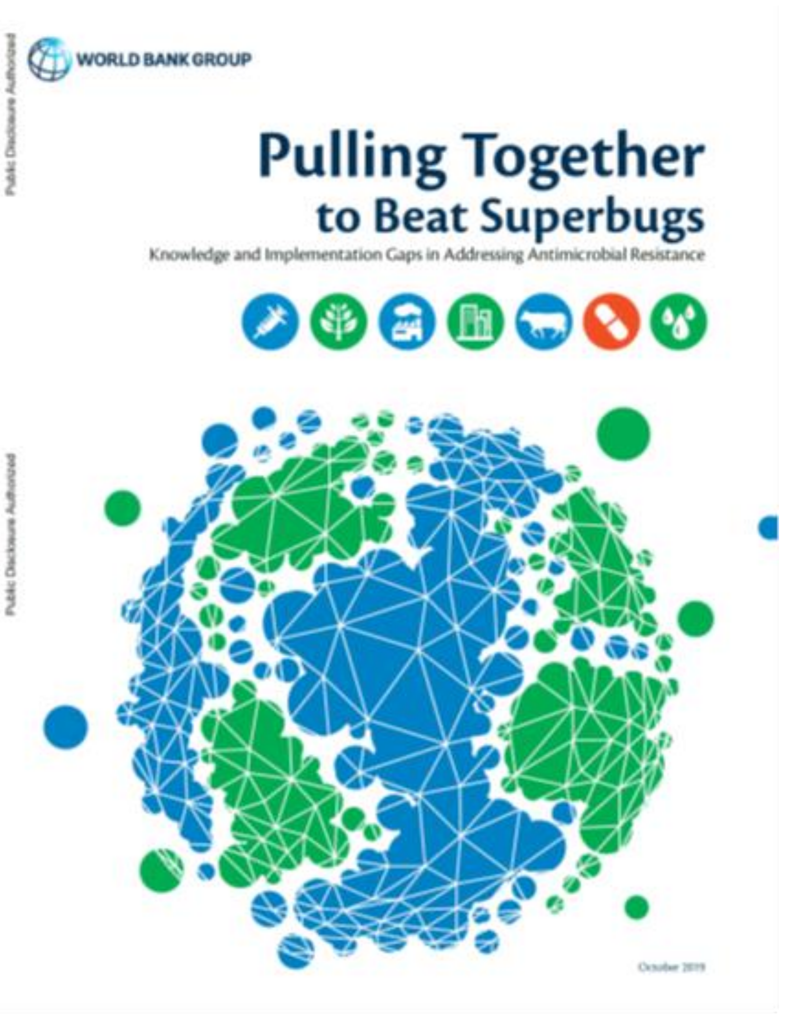
Predictors of higher AMR

- Open defecation
- Diarrhoea prevalence in children
- Mortality rate
- Death, by communicable diseases and maternal, prenatal and nutrition conditions
- Risk of maternal death
- Risk of impoverishing expenditure for surgical care
- Informal employment

Predictors of lower AMR

- Investment in water and sanitation
- Completeness of death reporting
- Educational attainment
- Number of surgical procedures
- Life expectancy at birth
- Number of Physicians
- Births attended by skilled health staff

Socio-economic parameters are more important than antimicrobial usage



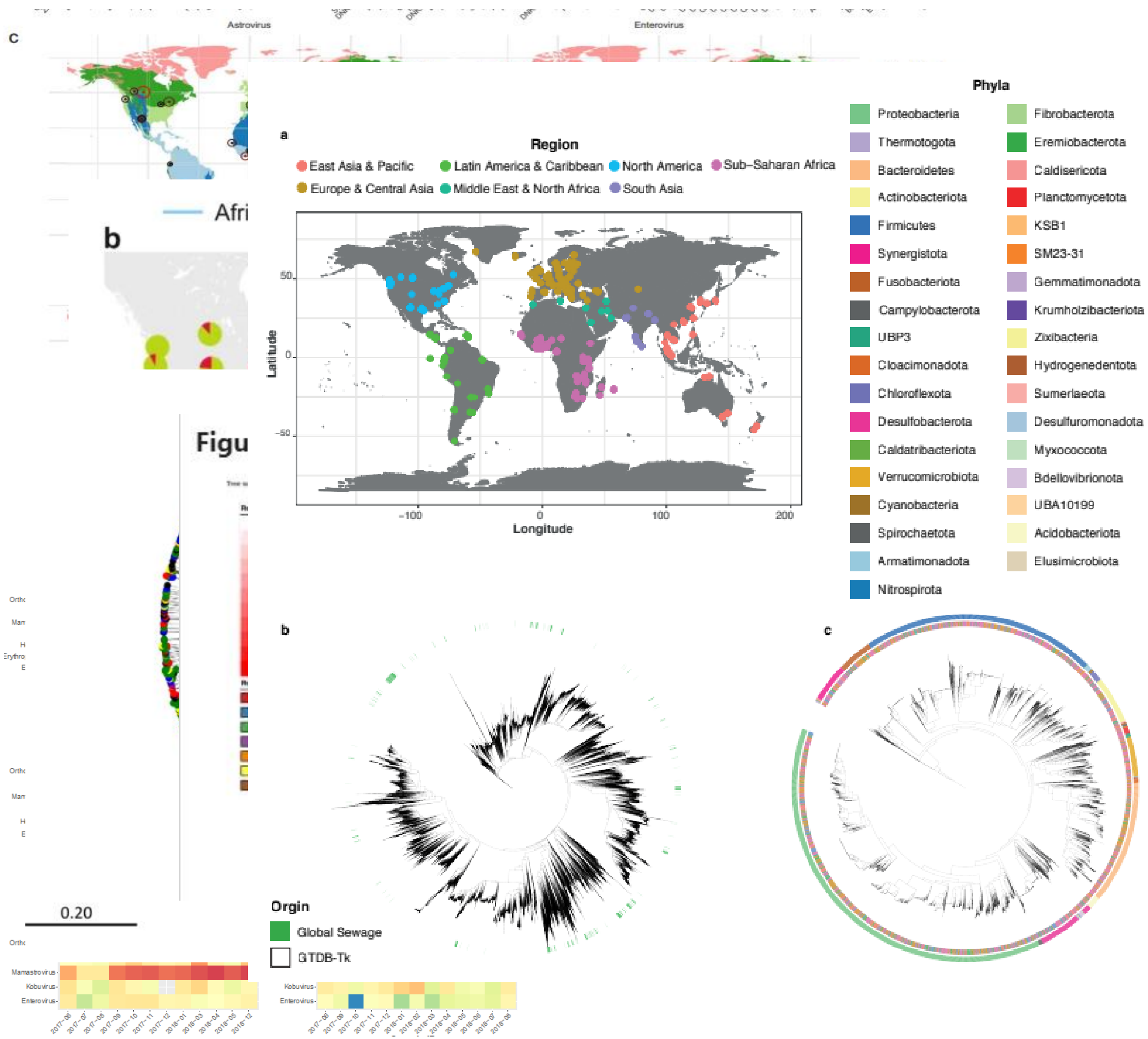
WORLD BANK GROUP

Pulling Together
to Beat Superbugs

Knowledge and Implementation Gaps in Addressing Antimicrobial Resistance

October 2019

Surveillance using wastewater – virus and everything else



Virus
Nieuwenhuijse et al. Sci. Rep. 2020.
Worp et al. in prep.

Human populations
Pipek et al. Sci. Rep. 2019

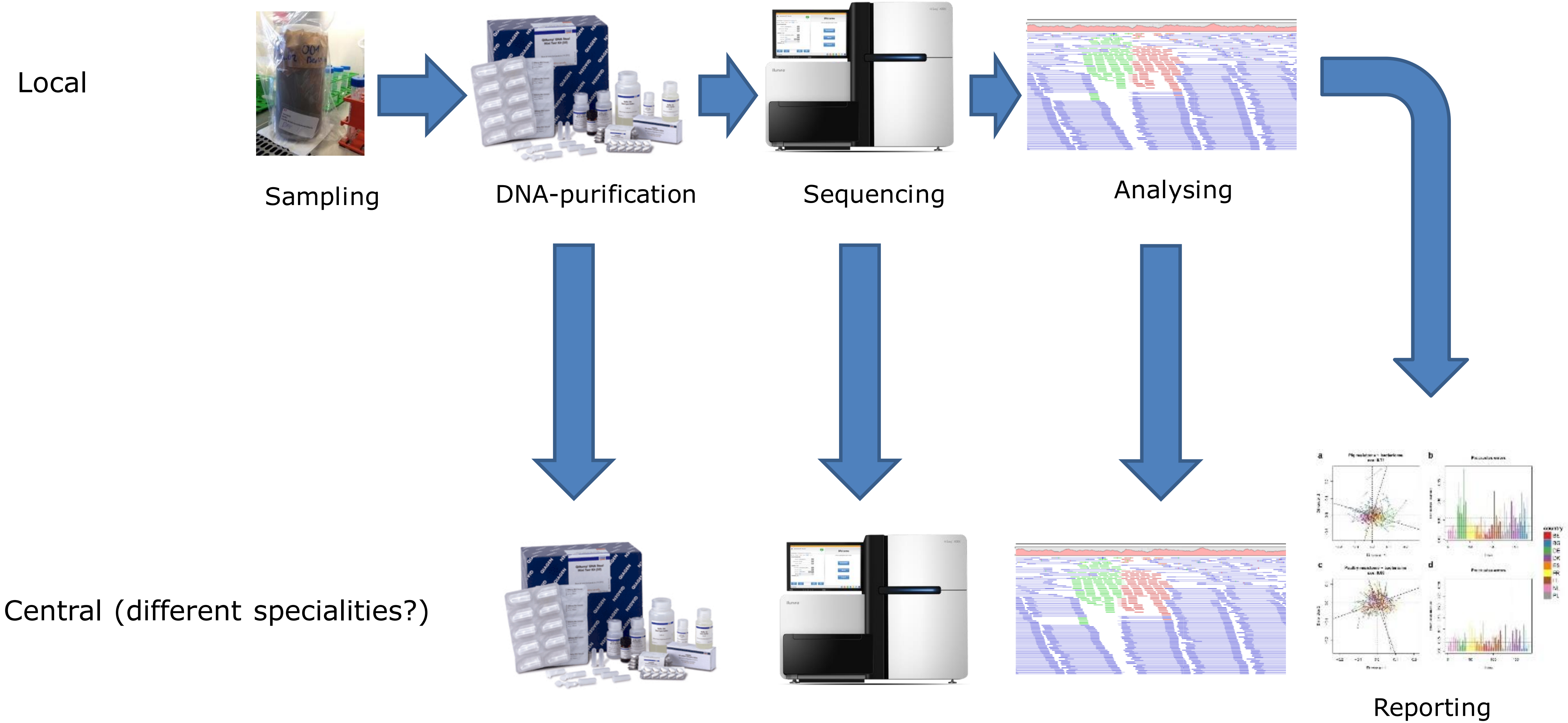
Plasmids
Teudt et al. mSystems. 2022

Bacteriophages
Edwards et al. Nature Microbiol. 2019

Bacterial phylogeny
Ahrenfeldt et al. Sci. Rep. 2020
Jespersen et al. in revision

How to create / implement global (human, animal or sewage) surveillance

Is it time to consider centralization / specialization?



Local

Sampling

DNA-purification

Sequencing

Analysing

Central (different specialities?)

Reporting

Other options

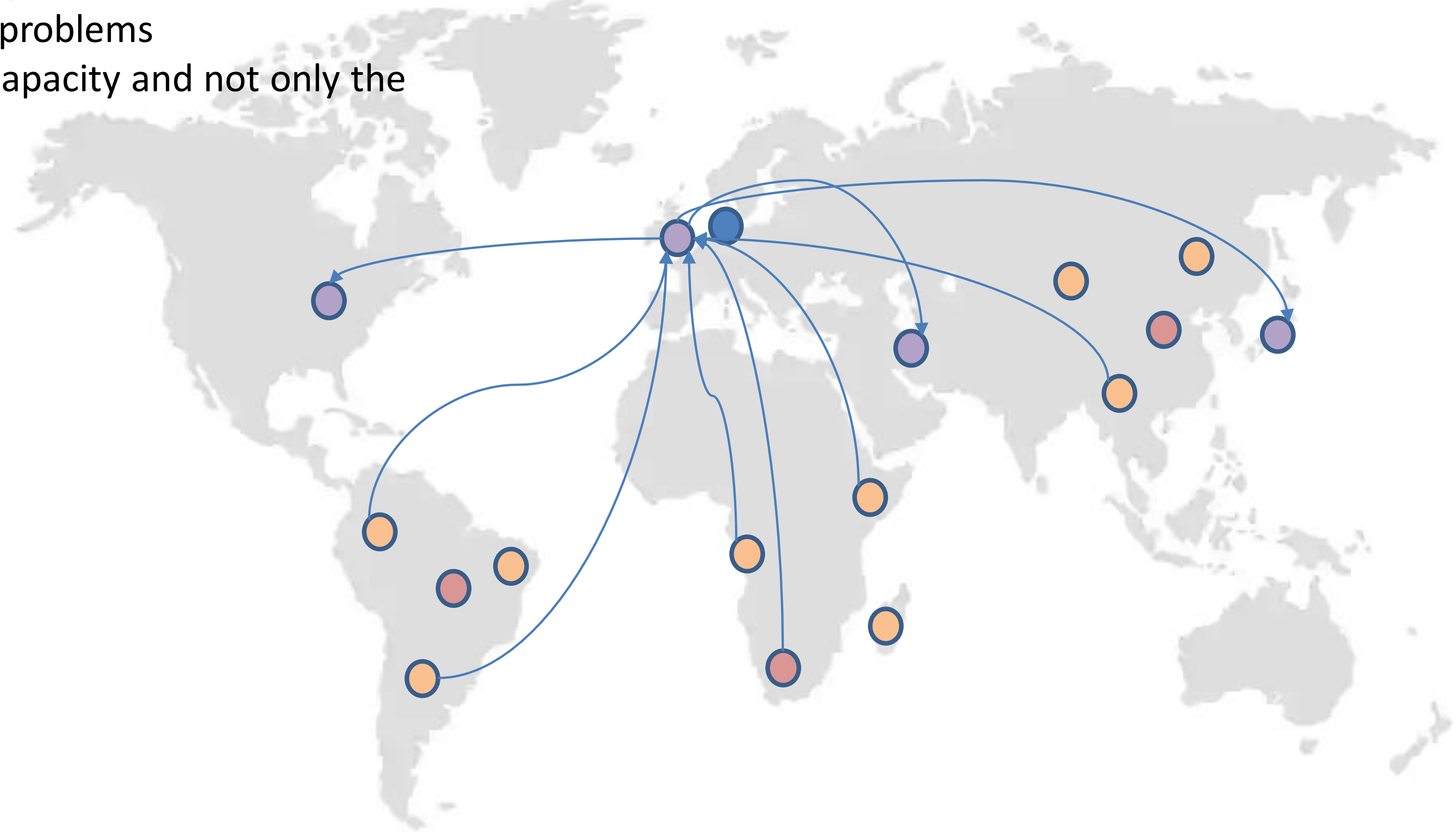
- **Perform**
  - Simply collect and analyse samples / isolates from the globe. Can be combined with training and capacity building

- **Enable**
  - Online tools
  - Frontline diagnostics and easy data sharing
  - Central private clouds (facebook of genomics)

- **Predict**
  - Improve data sharing options and utilize the randomly available data and build automatic models

Open science

- Allow analyses for all at a central facility
- Facilitating sharing with others and not only pre-defined problems
- Create frontline capacity and not only the big NRLs



2009 – online bioinformatics for dummies

Needs /possibilities

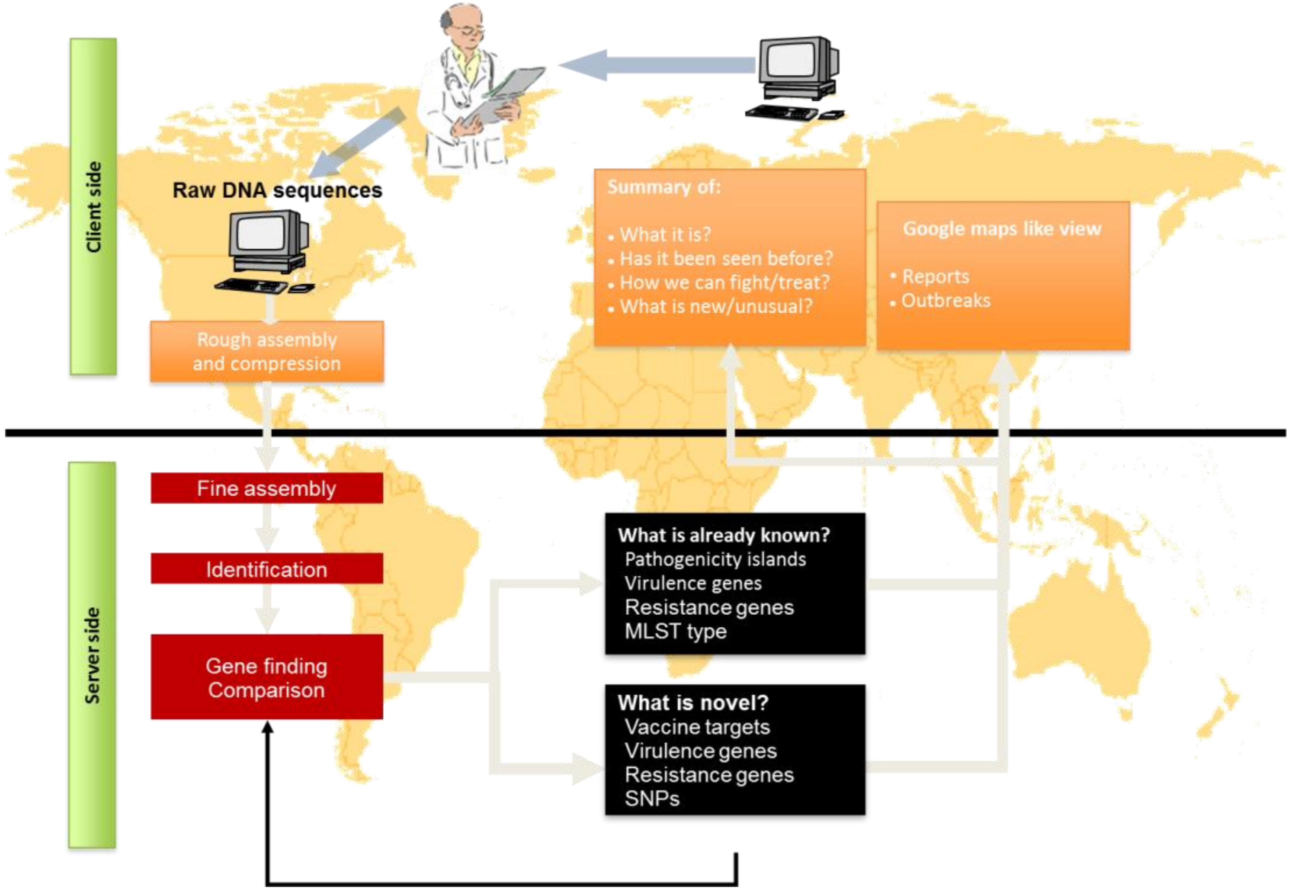
- More rapid response on global infectious disease transmission and sharing of data

Potential solution

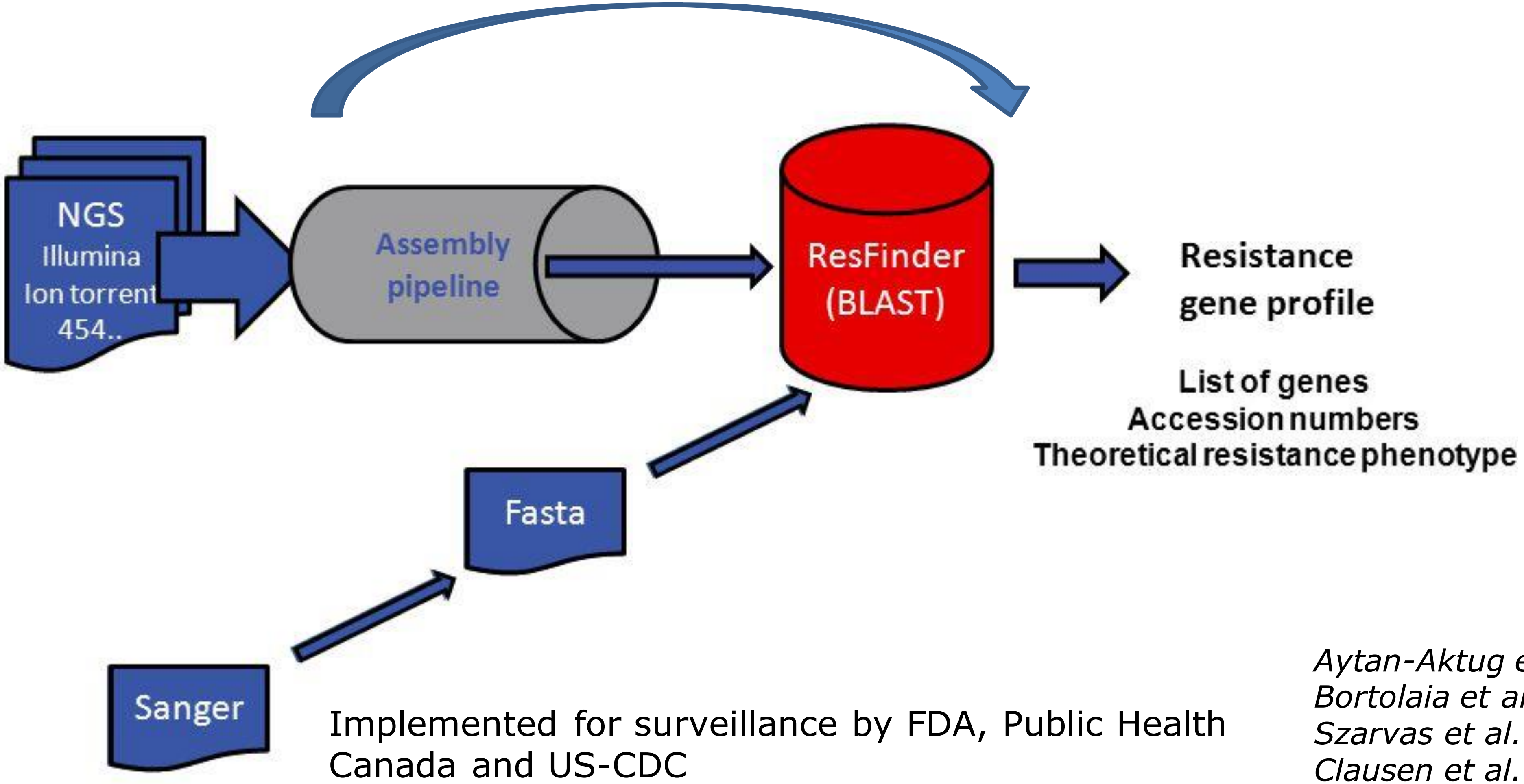
- NGS is a common language and easily shared

Development

- Provide online bioinformatics solutions



ResFinder



NGS
Illumina
Ion torrent
454..

Assembly pipeline

ResFinder (BLAST)

Resistance gene profile

List of genes
Accession numbers
Theoretical resistance phenotype

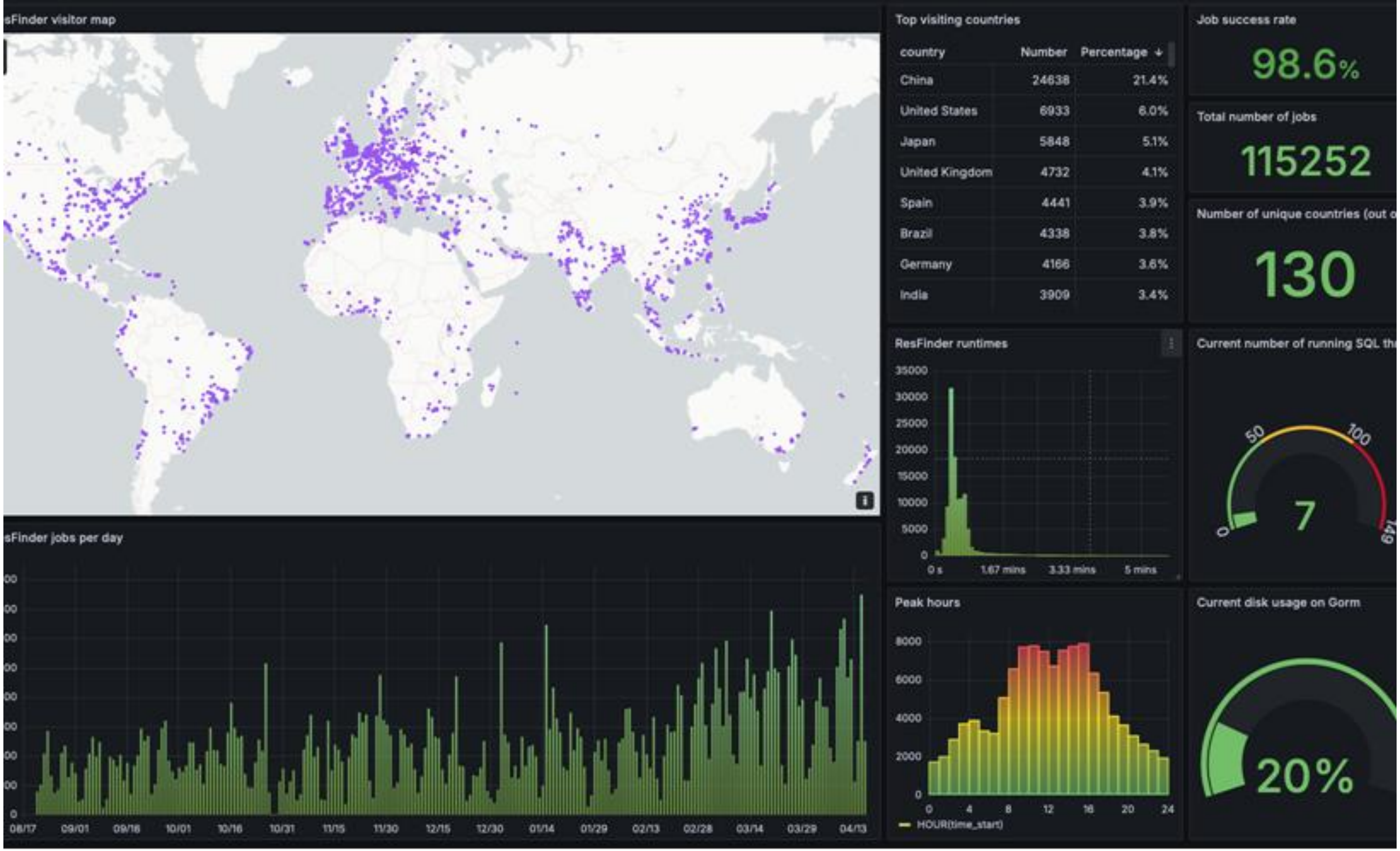
Fasta

Sanger

Implemented for surveillance by FDA, Public Health Canada and US-CDC

ResFinder 4.0 provides predicted phenotypes

Aytan-Aktug et al. mSystems. 2022
Bortolaia et al. JAC. 2020
Szarvas et al. Commun Biol. 2020
Clausen et al. BMC Bioinformatics. 2018
Kaya et al. mSphere. 2017
Zankari et al. JAC. 2017
Kaas et al. PlosOne. 2014
Zankari et al. JAC. 2012
Larsen et al. J Clin Microbiol. 2012



Top visiting countries

country	Number	Percentage ↓
China	24638	21.4%
United States	6933	6.0%
Japan	5848	5.1%
United Kingdom	4732	4.1%
Spain	4441	3.9%
Brazil	4338	3.8%
Germany	4166	3.6%
India	3909	3.4%

Job success rate
98.6%

Total number of jobs
115252

Number of unique countries (out o
130

Current number of running SQL th
7

Current disk usage on Gorm
20%

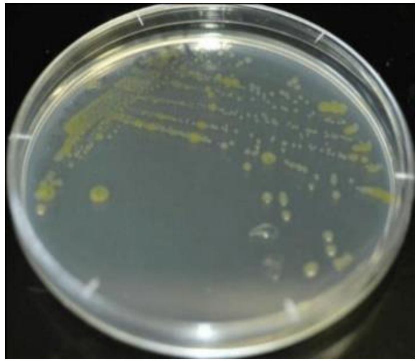
Current use is around 20,000 analyses per month for Resfinder and approx. the same for the other tools

Florensa et al. ResFinder - an open online resource for identification of antimicrobial resistance genes in next-generation sequencing data and prediction of phenotypes from genotypes. Microb Genom. 2022;8(1):000748.

Enabling the frontline

Pathogen sampling from patient and cultivation

Over night

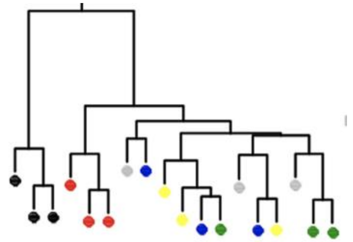


Purification And Library prep.

2 - 8 hours

Sequencing

0.5 - 3 hours



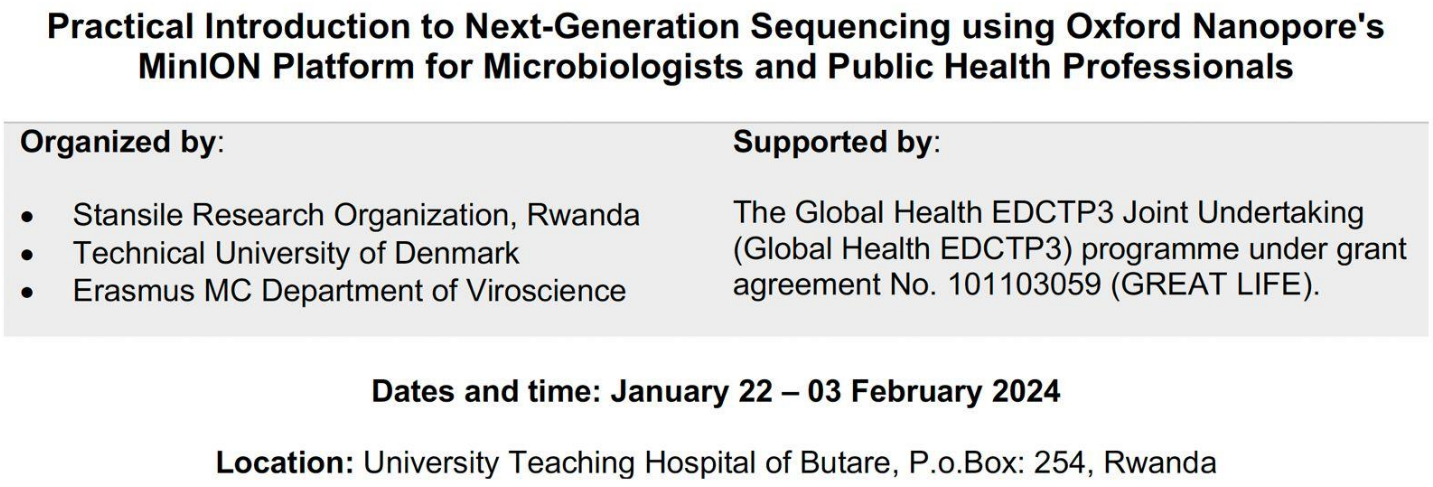
Analysis ~5min

- Bentolab purification
- ONT-sequencing
- Laptop analyses – automatic output
- Forced entering of meta-data for initiating analyses

Hallgren. Biol Methods Protoc. 2021

Kumburu et al. submitted

Kuchaka et al. submitted



Practical Introduction to Next-Generation Sequencing using Oxford Nanopore's MinION Platform for Microbiologists and Public Health Professionals

Organized by:

- Stansile Research Organization, Rwanda
- Technical University of Denmark
- Erasmus MC Department of Viroscience

Supported by:

The Global Health EDCTP3 Joint Undertaking (Global Health EDCTP3) programme under grant agreement No. 101103059 (GREAT LIFE).

Dates and time: January 22 – 03 February 2024

Location: University Teaching Hospital of Butare, P.o.Box: 254, Rwanda

Bas Oude Munnik

Wednesday 9.50

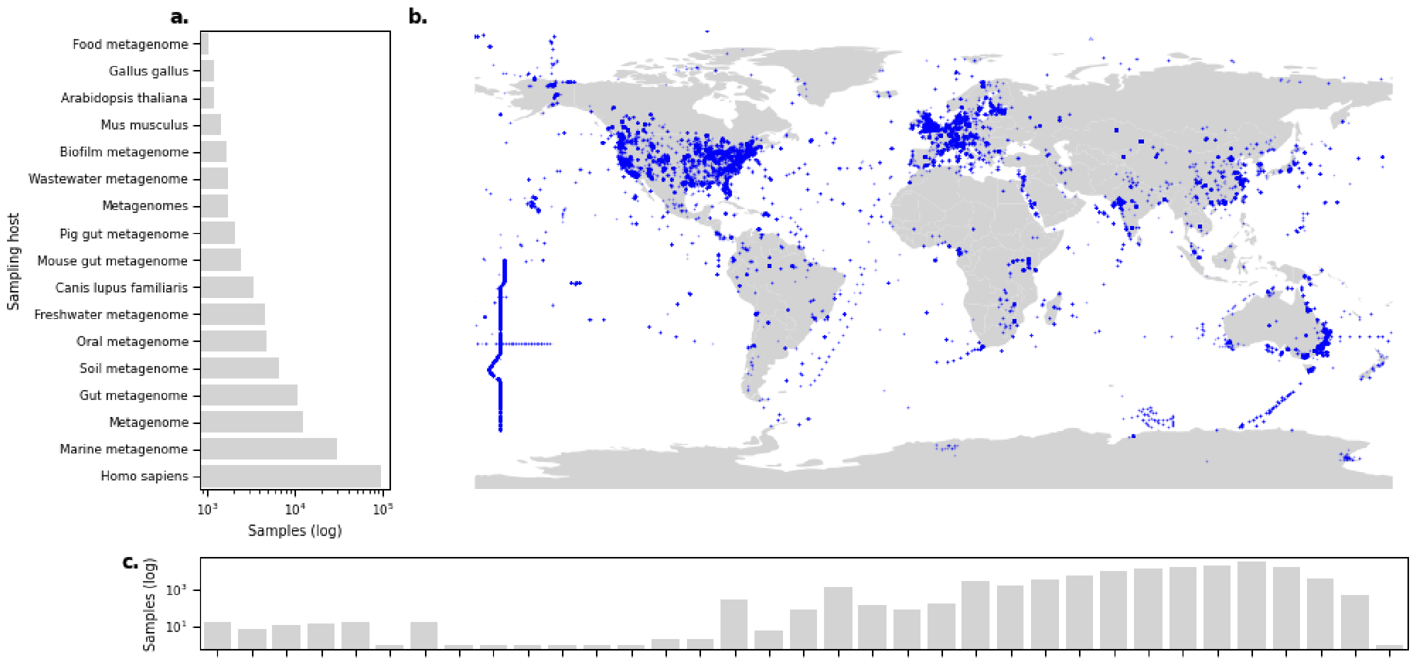
for more details

Other options

- **Perform**
  - Simply collect and analyse samples / isolates from the globe. Can be combined with training and capacity building

- **Enable**
  - Online tools
  - Frontline diagnostics and easy data sharing
  - Central private clouds (facebook of genomics)

- **Predict**
  - Improve data sharing options and utilize the randomly available data and build automatic models



Martiny et al. Plos Biol. 2022.

Martiny et al. mSystems. 2021

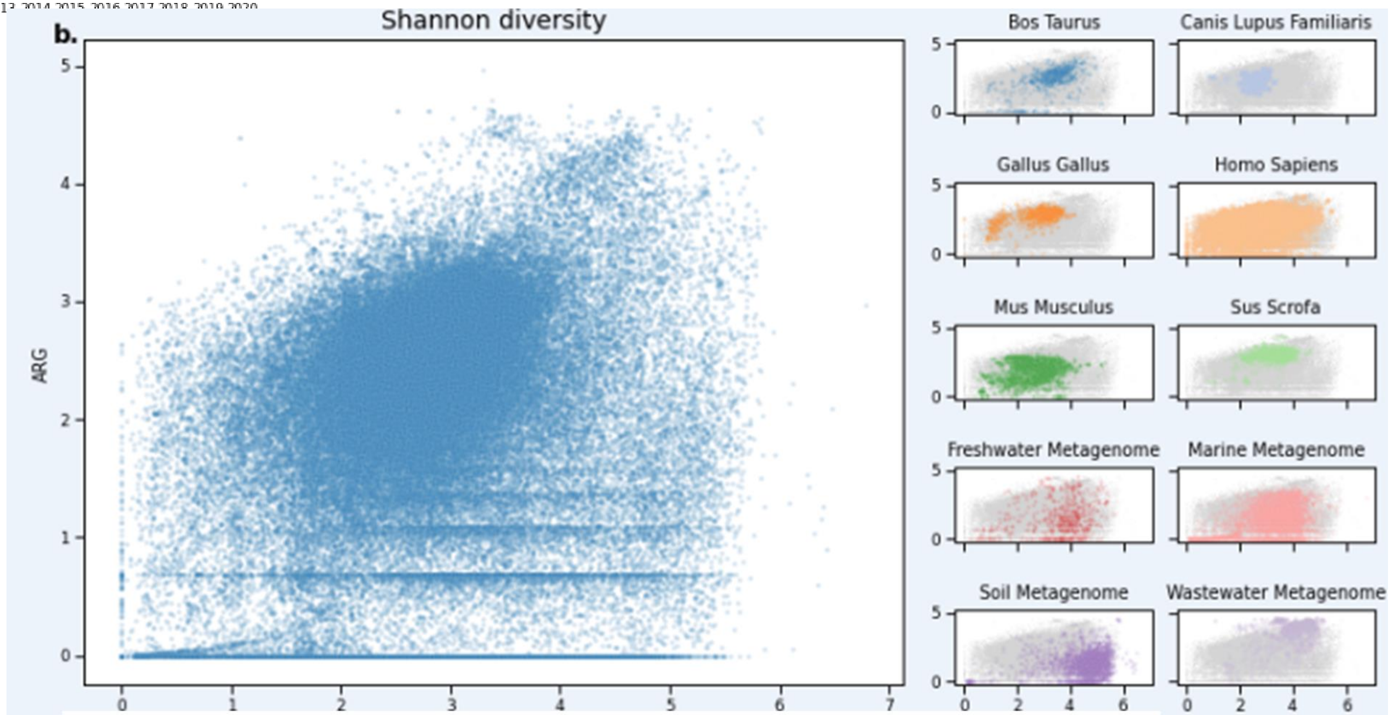
Analyse everything:

Download of all available metagenomic data

214,095 metagenomics datasets - 442 Terabytes - Download time 3 months

Mapped to all known ARGs - 2 months on a 40,000 core HPC

50 Gb table



Biodiversity is seemingly also relevant for One health AMR

Numbers – rough estimates

Activity	Rough estimated cost
The current WHO / ECDC / EFSA model	2-5 million € per country 400-1,000 million € per year
Centralised sewage or clinical surveillance – twice per year	1 million € per year
Airport surveillance (10 sites – each 25 samples per week)	3.5 million € per year
Online services for bioinformatics	250,000 € per year
Frontline sequencing, diagnostic and datasharing	10,000 € per site + 100,000 € for simple sharing hub per year
Public data accessibility	200,000 € per year

Other options

- **Perform**
  - Simply collect and analyse samples / isolates from the globe. Can be combined with training and capacity building

- **Enable**
  - Online tools
  - Frontline diagnostics and easy data sharing
  - **Central private clouds (facebook of genomics)**

- **Predict**
  - Improve data sharing options and utilize the randomly available data and build automatic models

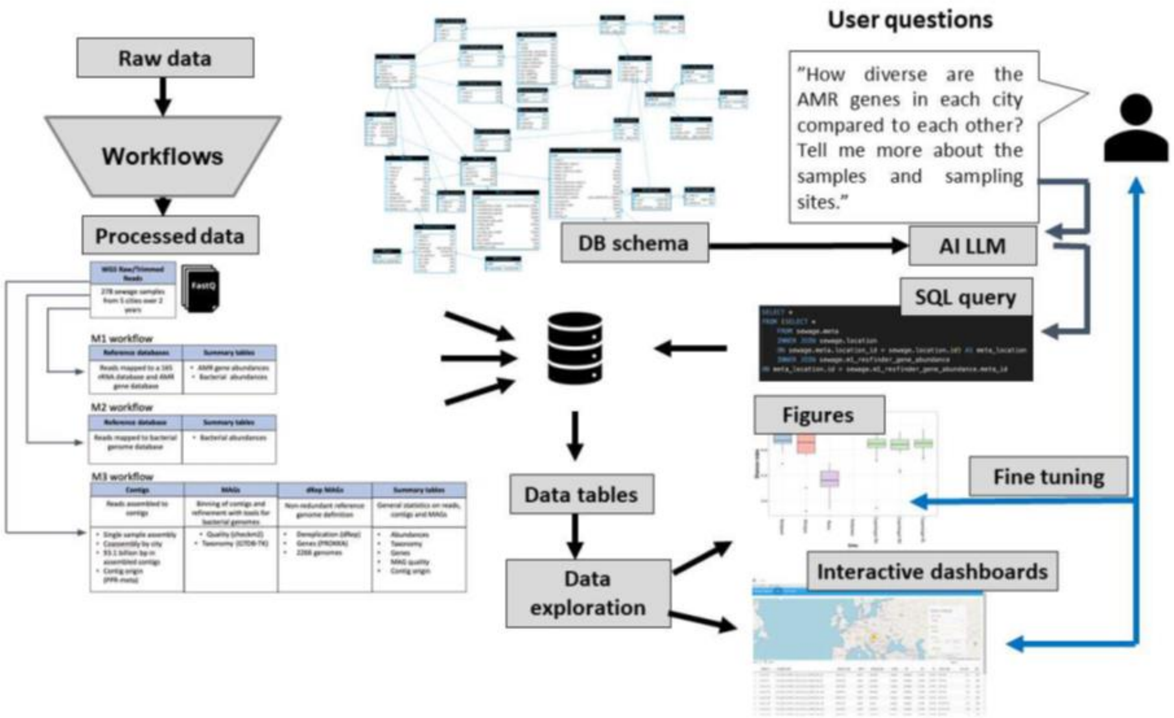
So where to go?

- The current WHO / ECDC / EFSA / NRL system is extremely expensive and not always (sufficient) flexible or integrative
- The need to control “Own” data locally is a major driver
- There is a need to:
  - Enable the frontline and active individuals
    - Online tools
    - Better and simpler lab-top solutions
    - We need a flexible datasharing site (facebook of genomics)

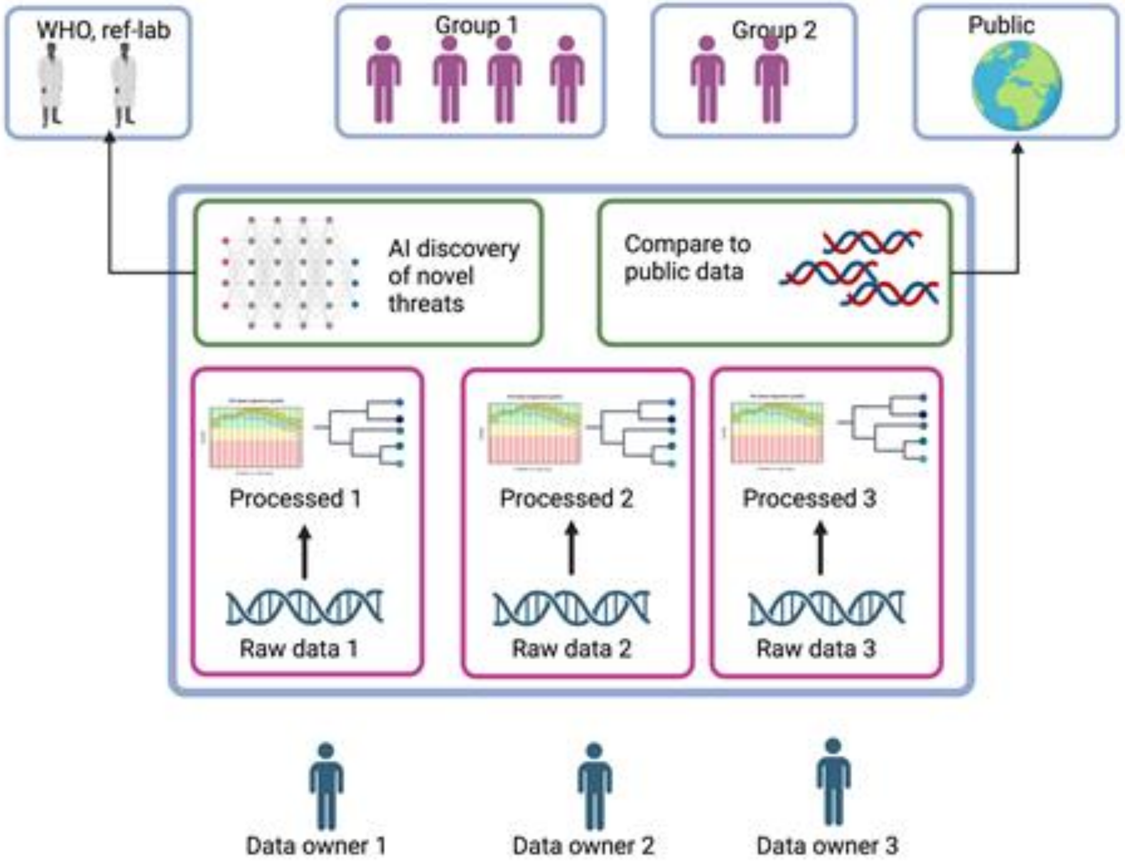
Leveraging the complexity of data

- Big data are complex
- Major need for (bio)informatics and organization
- And ability to share in controlled groups

- Different solutions tested in VEO



"How diverse are the AMR genes in each city compared to each other? Tell me more about the samples and sampling sites."



A “cloud” solution for:

- Individual upload and analyses of own data
- Complete control of sharing what with whom
- Possible to compare to public data
- Possible to create collaborative groups
- Potentially possibilities to do automatic screening for novel signal without accessing the data

- Estimated cost: 500 million € to establish and 20 million € in annual running cost

- **I take cash**

Concluding remarks

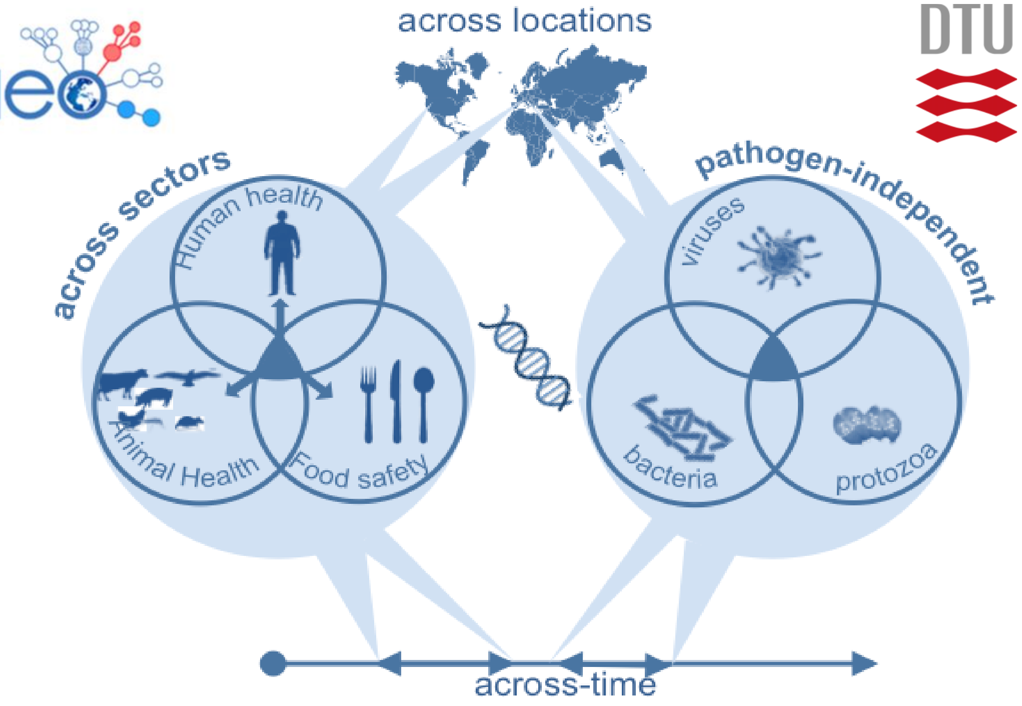
- Surveillance is the basis of everything, and without it we are flying blind
  - Easy to pilot – surprisingly difficult to conduct (sewage, patients, perhaps public data?)
  - Perhaps time to rethink and not wait until all countries can do everything?

- (raw) Data-sharing and complete transparency is essential and should be a mandatory part of any funding

- Prevention is better than cure (but fame and fortune is in the cure)

  (Bian Que 407-310 BC)

Our vision: one system serves all



Guiding principles:

- Cross sector, cross domain, open source (not commercial)
- Interaction with the rest of the world (all inclusive)
- Data for action (actionable outputs)
- Central repository (ENA, DDJ, NCBI)

The Danish Council for Strategic Research
VILLUM FONDEN
DANIDA
Ministry of Foreign Affairs of Denmark
UKaid
from the British people
The Fleming Fund
HORIZON 2020
European Commission
novo nordisk fonden