Antigen Surveillance: from Evolution to Immune Escape

Streptococcus pneumoniae: vaccine strategies and antigenic variation

Dr Stephanie Lo
Team Leader
EMBL's European Bioinformatics Institute

Intended Learning Objectives

Specific objectives of this session:

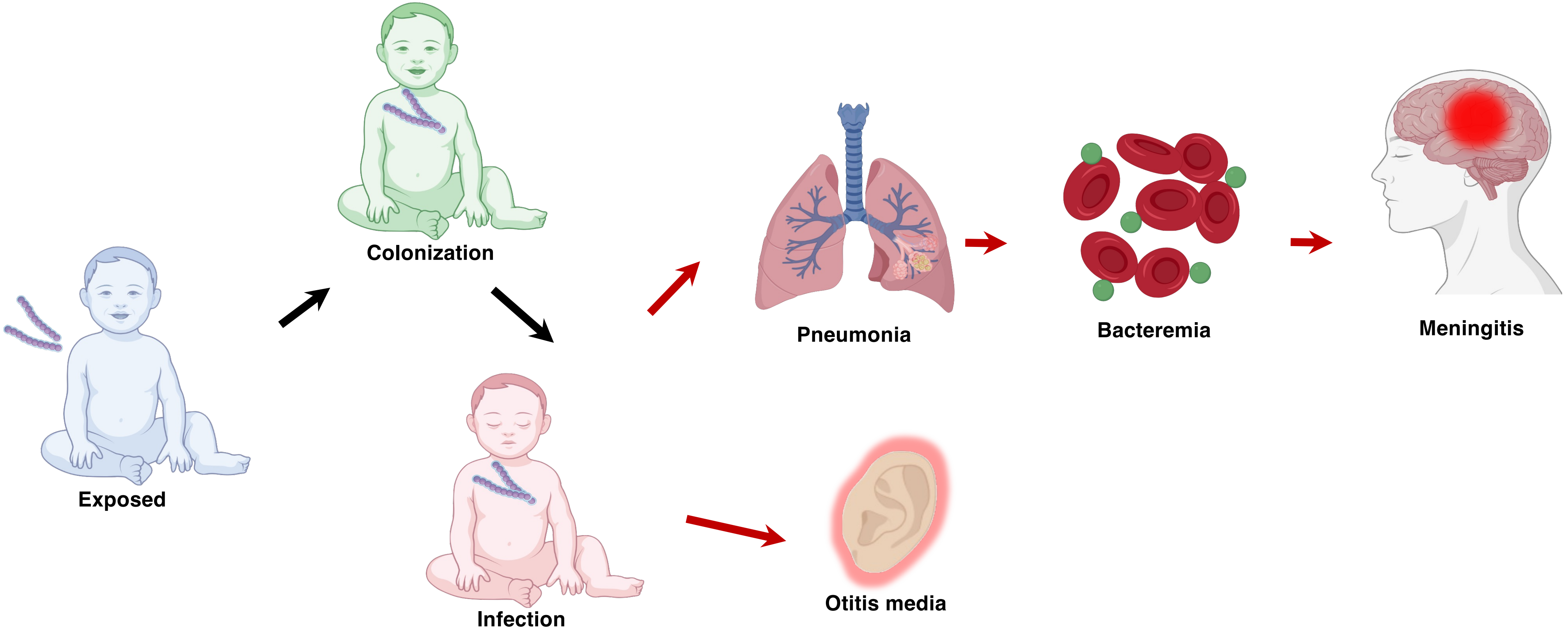
1. Biodiversity and population structure
2. Typing methods
3. Genomic surveillance to evaluate vaccine impact
4. Track vaccine-evading strains

Outline

This session consists of the following elements

1. What genomic typing method could be used to define pneumococcal population structure?
2. How pneumococci evade the vaccine?
3. What is the added value of genomics in pathogen surveillance?

Streptococcus pneumoniae



Exposed

Colonization

Infection

Pneumonia

Bacteremia

Meningitis

Otitis media

Created in BioRender.com bio

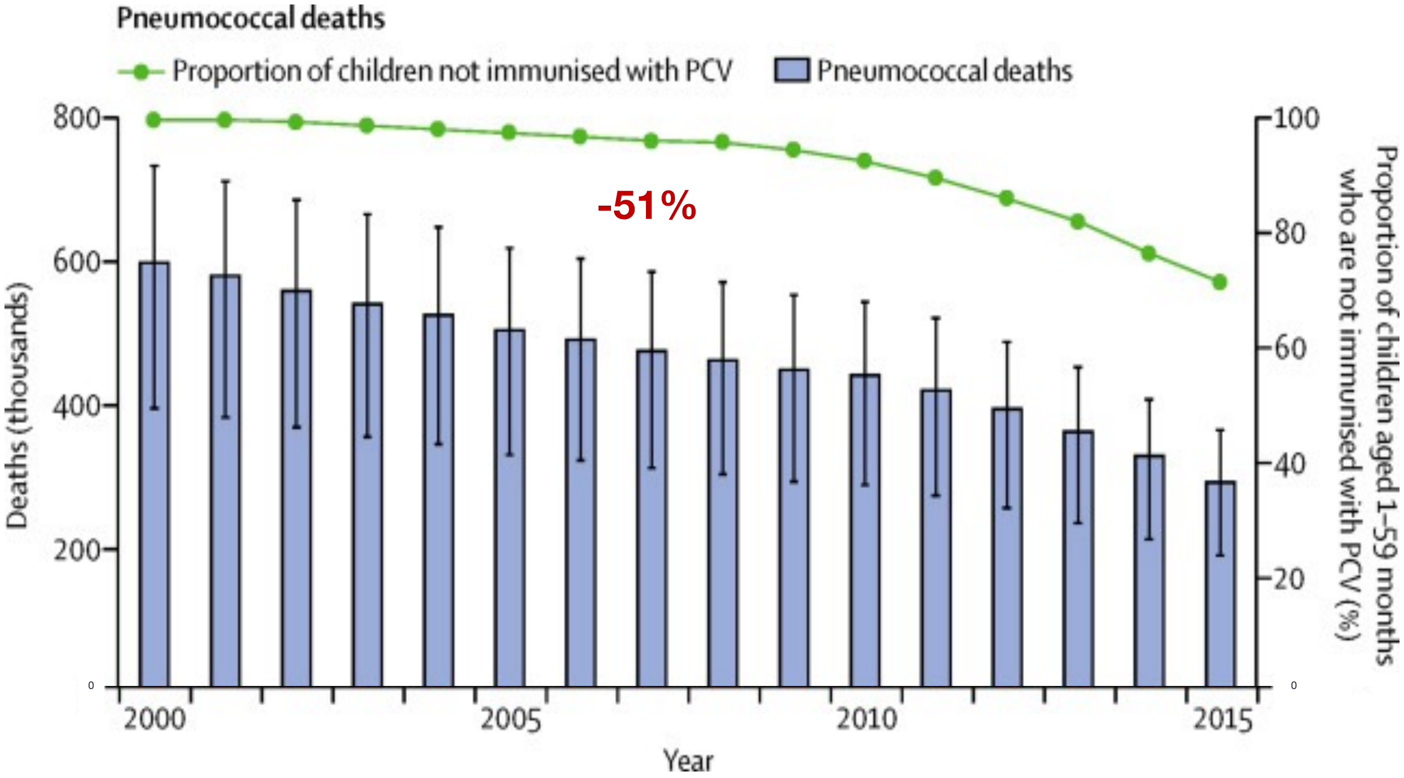
Pneumococcal Conjugate Vaccine

107 serotypes

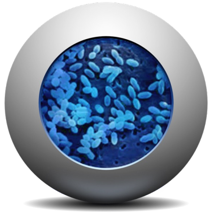
Kim et al 1999 I&I

PCV7, 10, 13, 15, 20

https://www.europeanpharmaceuticalreview.com/



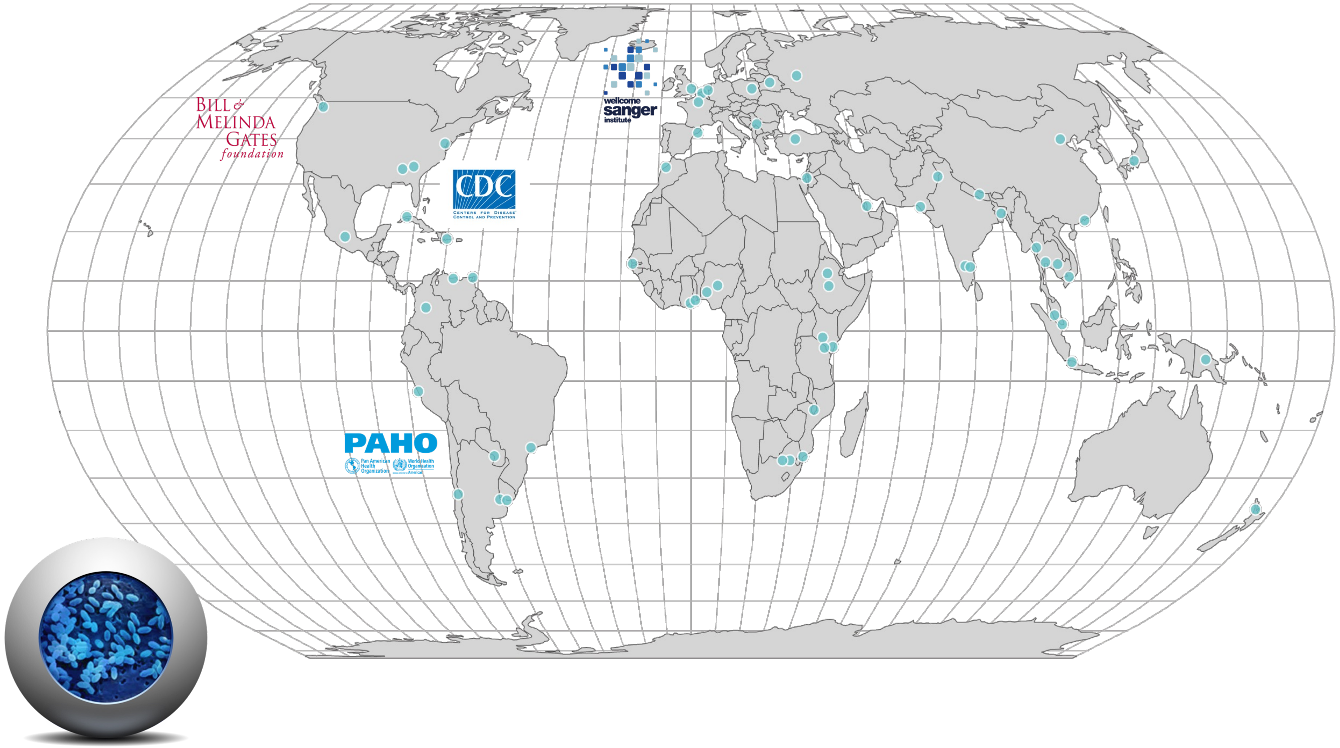
Wahl et al 2018 Lancet Glob Health



Global **P**neumococcal **S**equencing Project

2011- present

Global Pneumococcal Sequencing (GPS) project



>40,000 genomes

>60 countries

1989-2024

Outline

This session consists of the following elements

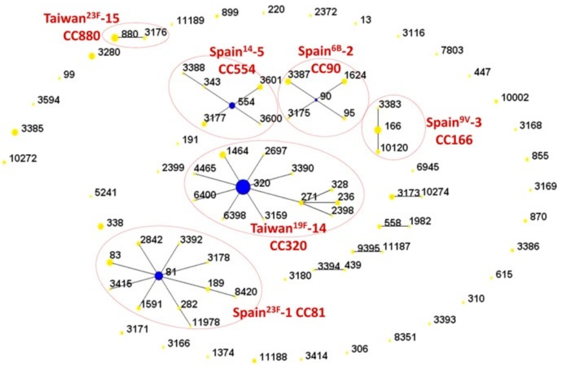
1. What genomic typing method could be used to define pneumococcal population structure?
2. How pneumococci evade the vaccine?
3. What is the added value of genomics in pathogen surveillance?

Do not edit
How to change the design

How would you characterize the population sturcture of over 40,000 pneumococcal genomes?

Defining pneumococcal lineages

MLST (ST and CC)

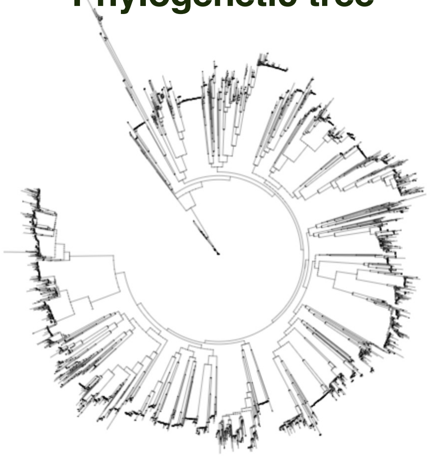


Yun et al 2018

Pros: standardised nomenclature

Cons: 7 genes resolution, **obscure by frequent recombination**

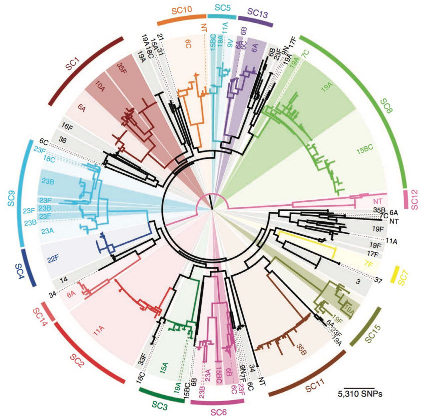
Phylogenetic tree



Pros: WGS resolution

Cons: large dataset, **subjective cut-off,** new samples unfriendly

HierBAPs



Croucher et al 2013

Pros: WGS resolution, objective clustering

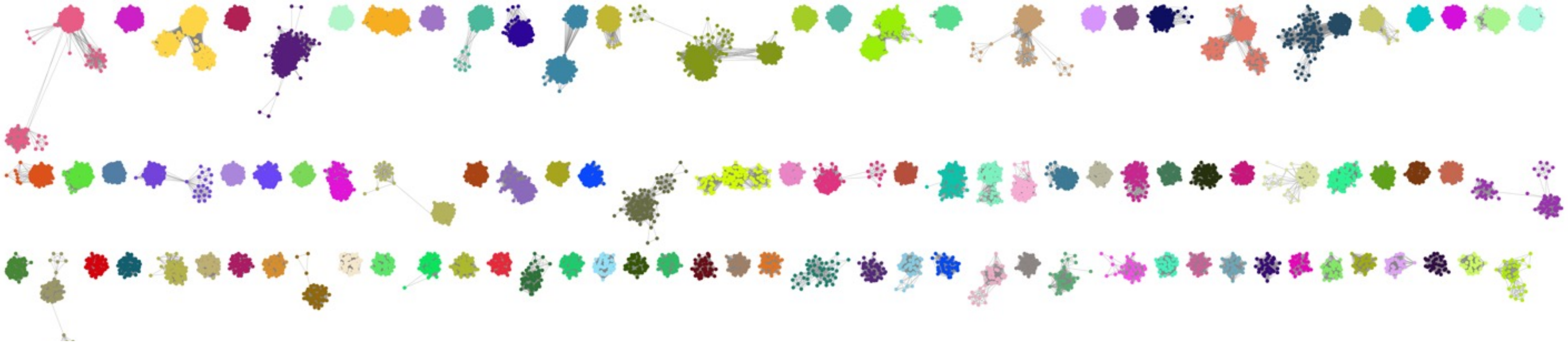
Cons: **No standardised nomenclature, new samples unfriendly, bin**

Global Pneumococcal Sequence Cluster (GPSC)

- Whole-genome resolution
- Cluster down to a single strain (no bin)
- Standardised nomenclature
- Reference database update

PopPUNK
PopPUNK
PopPUNK
PopPUNK
PopPUNK

Lees and Croucher et al 2019 Genome Research

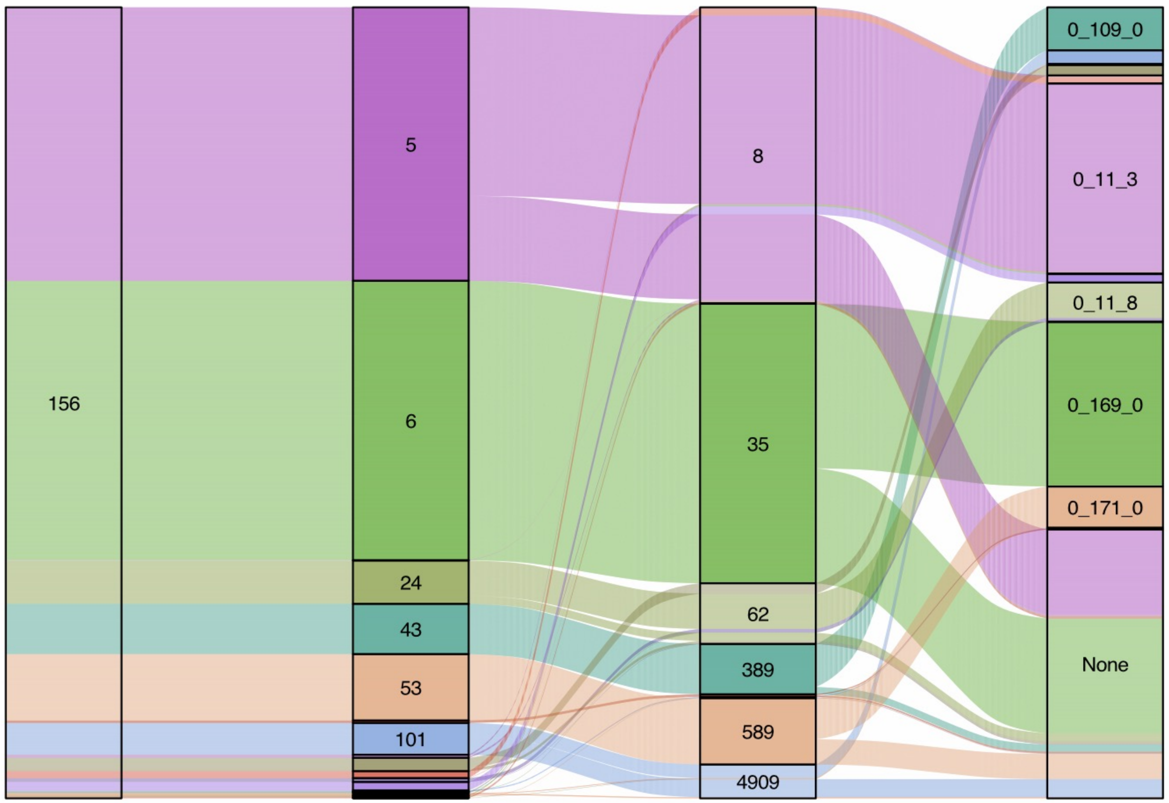


Ref database v8

- 42,493 genomes
- 977 GPSCs

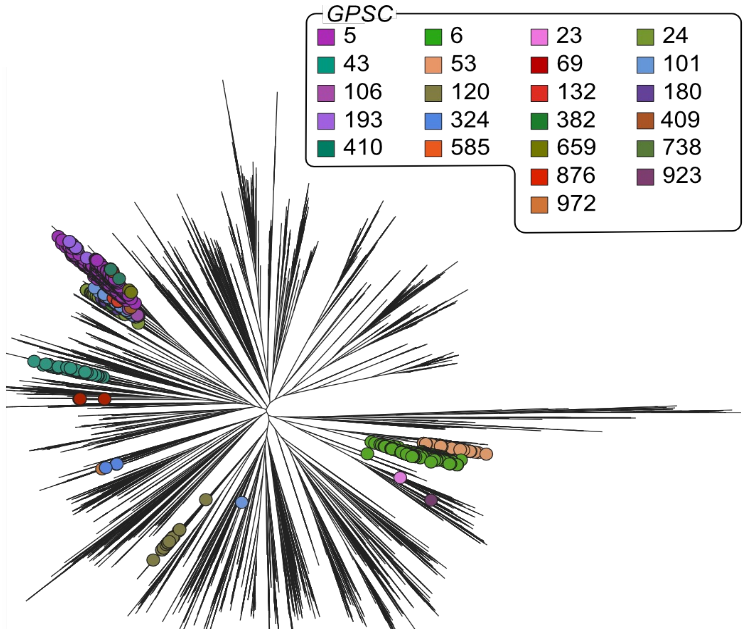
Gladstone and Lo et al 2019 EbioMedicine

GPSC vs CC vs cgMLST vs LIN code



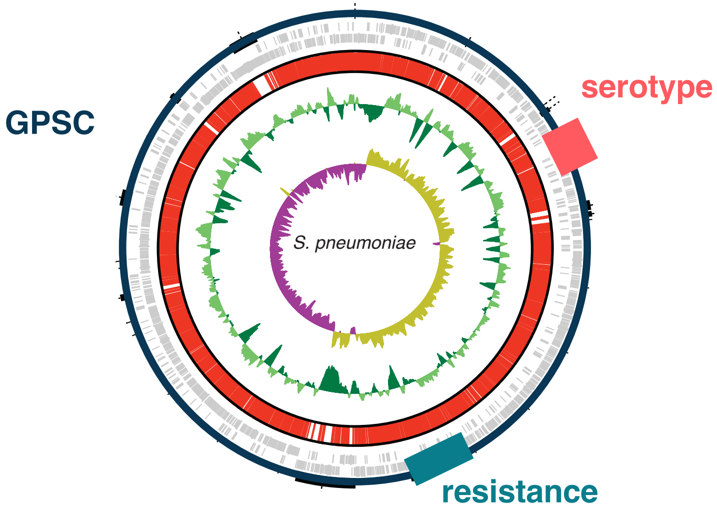
CC | GPSC | cgMLST | LIN code

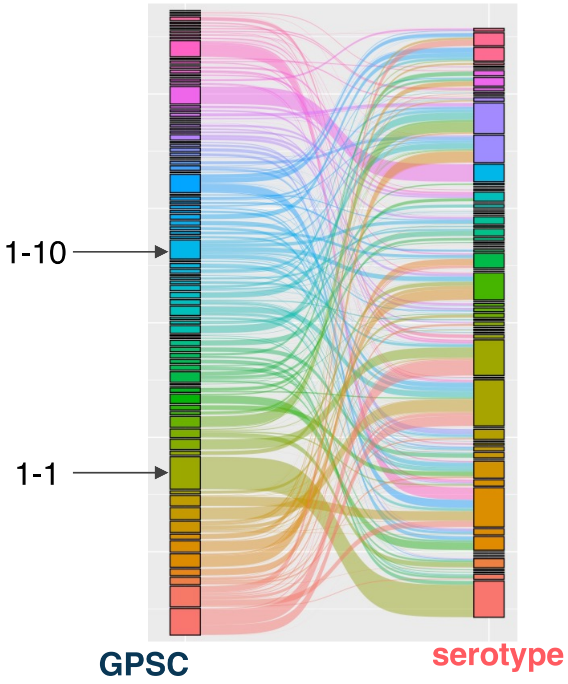
7-gene | whole-genome | Core genes



King et al 2024 Microb Genom

GPSC vs serotyping





Streptococcus pneumoniae is genetically diverse
107 serotypes, 977 lineages

What information can extract from a genome?

Analyse your pneumococcal genomes

using

Pathogen**watch**

An all-in-one bioinformatics software

GPS PIPELINE

- **Input**

Pair-end fastq files

- **Output 1 (.csv)**

✅ QC

✅ Pneumococcal lineage (GPSC)

✅ In silico serotype

✅ MLST

✅ Predict susceptibilities to 19 common antibiotics

Microreact

- **Output 2 (.csv)**

✅ software version

- **Output 3 (.fa & .gff)**

✅ de novo assemblies

✅ annotation

Hung et al 2024 Nat Comm
https://github.com/GlobalPneumoSeq/gps-pipeline

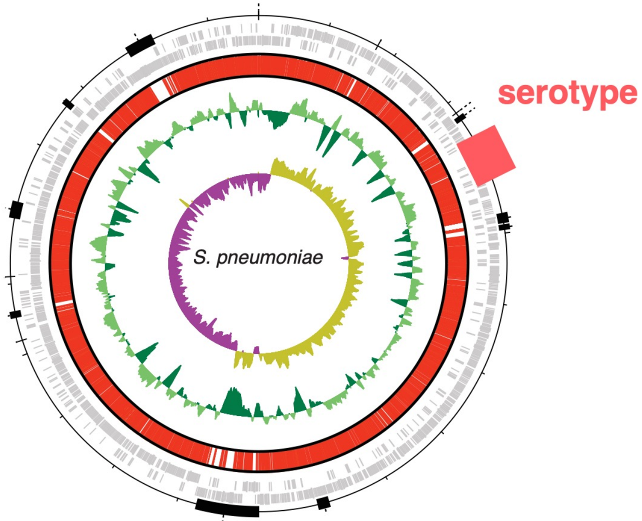
Outline

This session consists of the following elements

1. What genomic typing method could be used to define pneumococcal population structure?
2. How pneumococci evade the vaccine?
3. What is the added value of genomics in pathogen surveillance?

Vaccine evasion

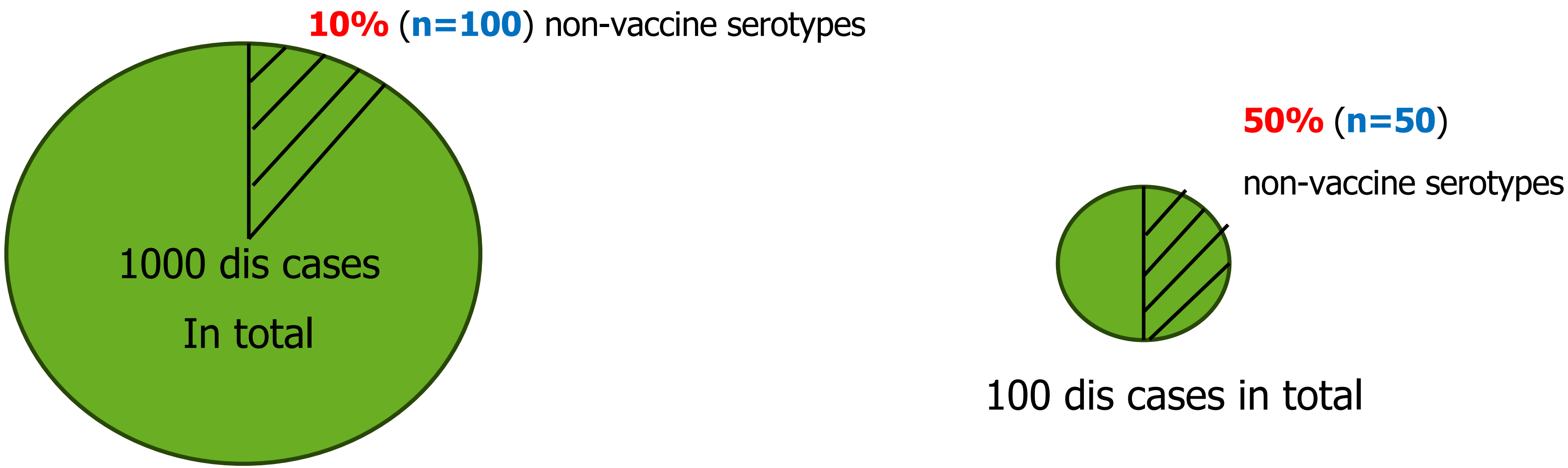
- Vaccine breakthrough or failure
- Serotype replacement



Serotype replacement

Serotype replacement is defined by a significant increase in disease cases (or incidence rate) caused by pneumococci that expressed serotypes not covered by the current vaccine

Why not in proportion/prevalence?



10% (**n=100**) non-vaccine serotypes

1000 dis cases

In total

Pre-vaccine

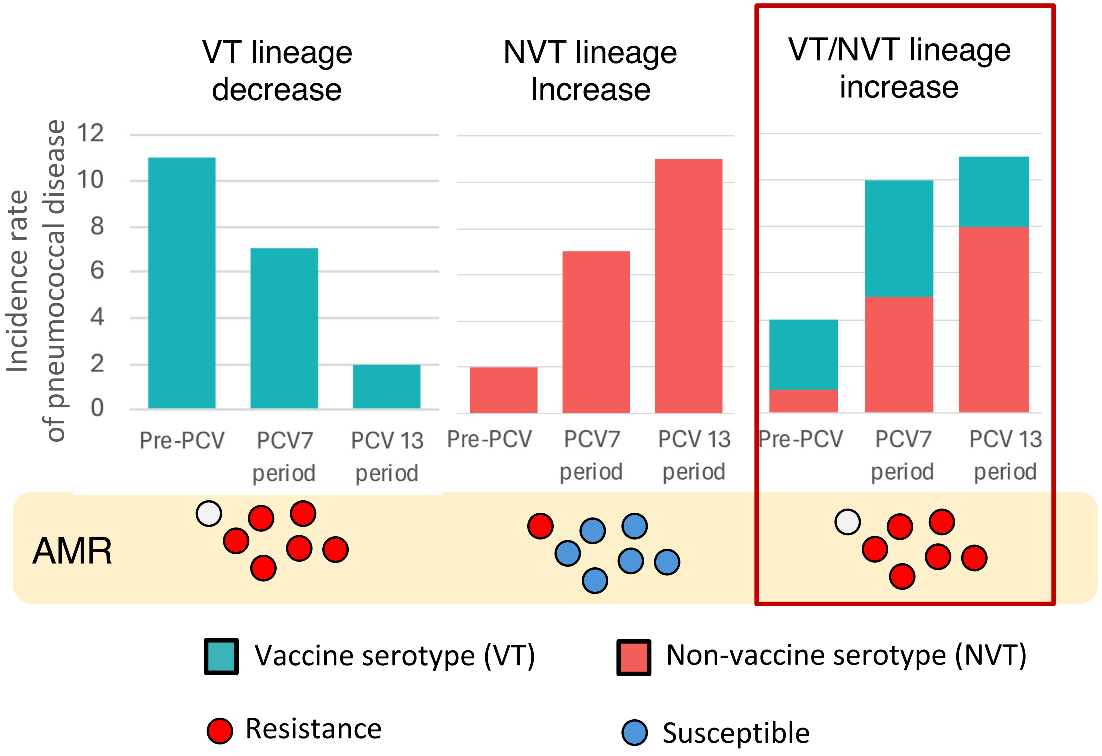
50% (**n=50**)

non-vaccine serotypes

100 dis cases in total

Post-vaccine

Vaccine-selective pressure



Incidence rate ratio
(post-PCV13 reference to pre-PCV)

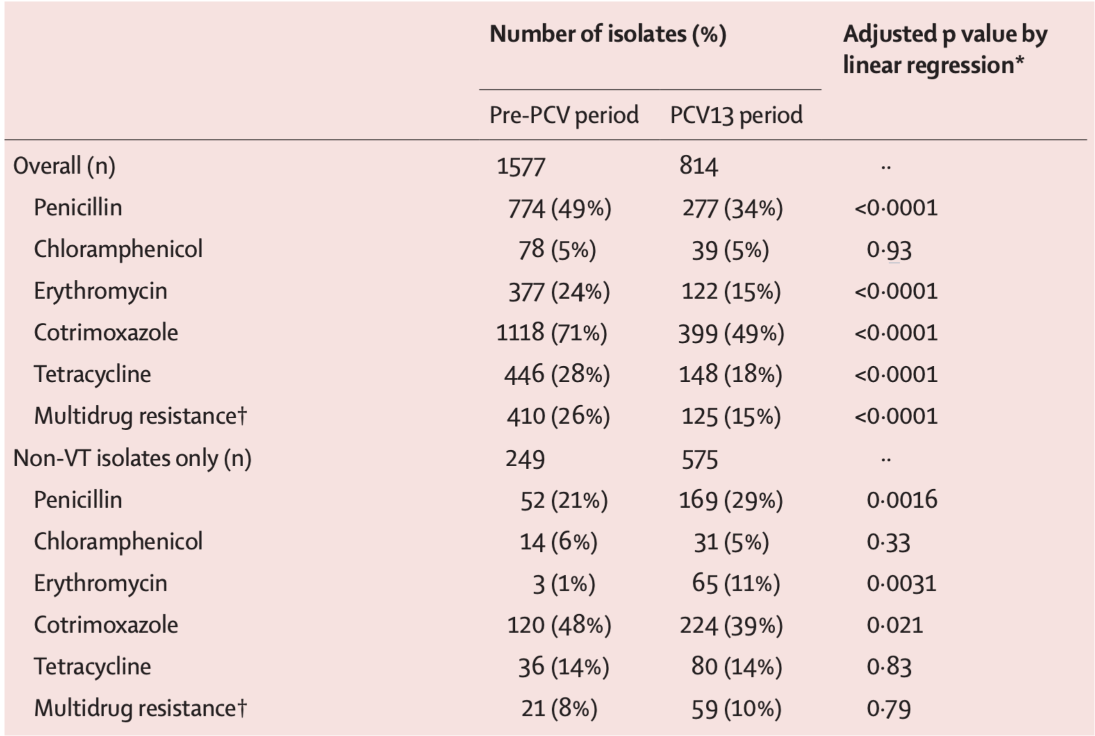
	NVT lineage	VT/NVT lineages
Israel	1.9 (1.4-2.6)*	3.4 (1.8-6.3)*
South Africa	1.1 (0.7-1.6)	2.5 (1.6-3.9)*
USA	0.9 (0.7-1.2)	1.6 (1.0-2.4)

*Statistically significant

Lo & Gladstone et al 2019 The Lancet Infect Dis

Does pneumococcal conjugate vaccine reduce AMR?

Pooled prevalence of Israel, South Africa and USA



	Number of isolates (%)		Adjusted p value by linear regression*
	Pre-PCV period	PCV13 period	
Overall (n)	1577	814	··
Penicillin	774 (49%)	277 (34%)	<0·0001
Chloramphenicol	78 (5%)	39 (5%)	0·93
Erythromycin	377 (24%)	122 (15%)	<0·0001
Cotrimoxazole	1118 (71%)	399 (49%)	<0·0001
Tetracycline	446 (28%)	148 (18%)	<0·0001
Multidrug resistance†	410 (26%)	125 (15%)	<0·0001
Non-VT isolates only (n)	249	575	··
Penicillin	52 (21%)	169 (29%)	0·0016
Chloramphenicol	14 (6%)	31 (5%)	0·33
Erythromycin	3 (1%)	65 (11%)	0·0031
Cotrimoxazole	120 (48%)	224 (39%)	0·021
Tetracycline	36 (14%)	80 (14%)	0·83
Multidrug resistance†	21 (8%)	59 (10%)	0·79

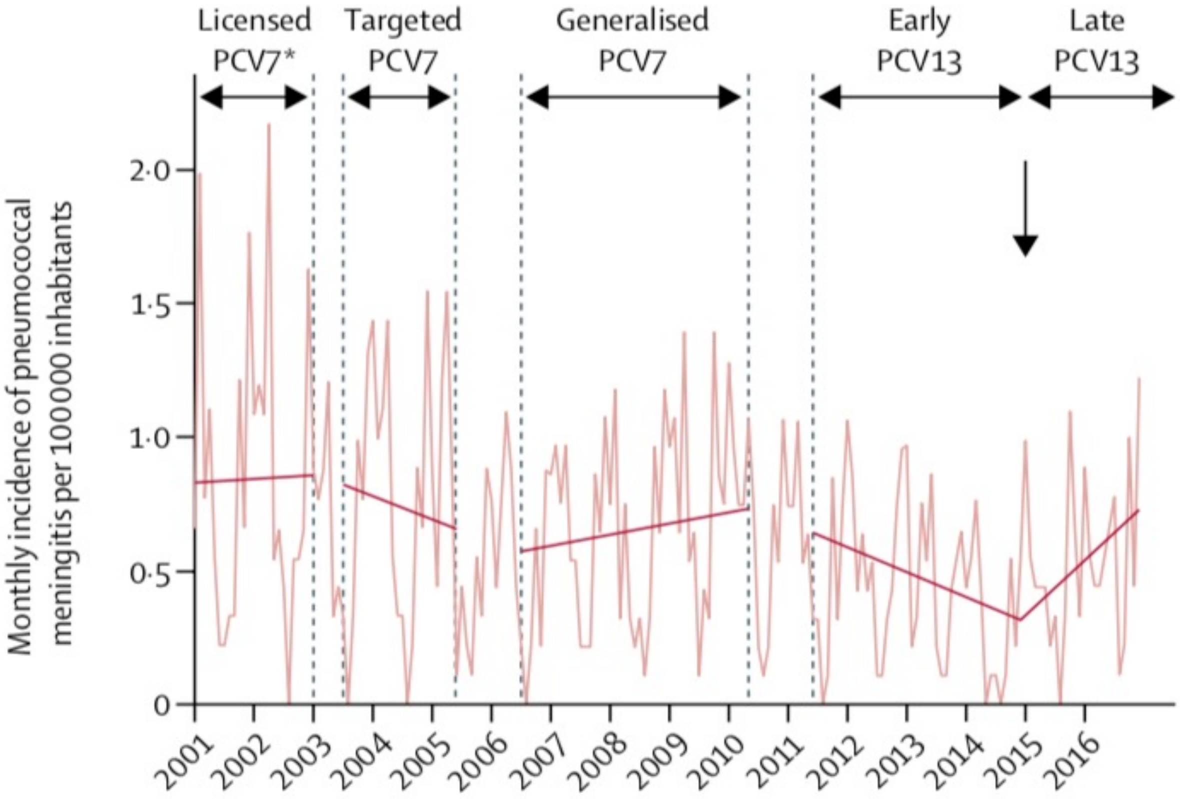
Lo & Gladstone et al 2019 The Lancet Infect Dis

Case study

Emerging serotype 24F in France

increase in penicillin-resistance in parallel

Meningitis cases from children <2 years



Ouldali et al 2018 The Lancet Infect Dis

What do we know about serotype 24F?

24F increased or became predominant

Argentina, Belgium, Denmark, Germany, Israel, Italy, Japan, Portugal, Lebanon, Norway, Spain, UK

High invasiveness

High invasive disease potential

Prone to cause meningitis

(Balsells et al 2018, meta-analysis)

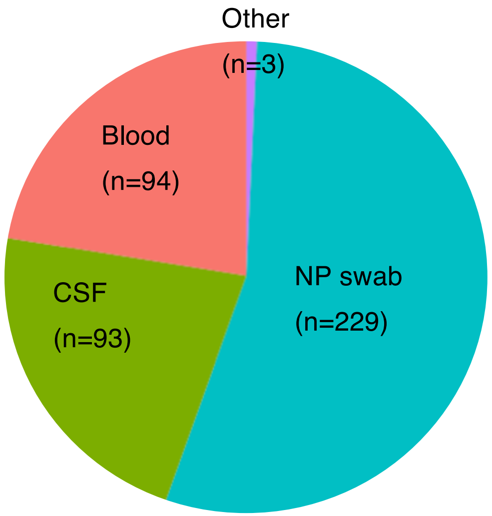
Prevention

NOT in current highest valent PCV20

NOT in upcoming PCV24

But in PCV21 (adult use)

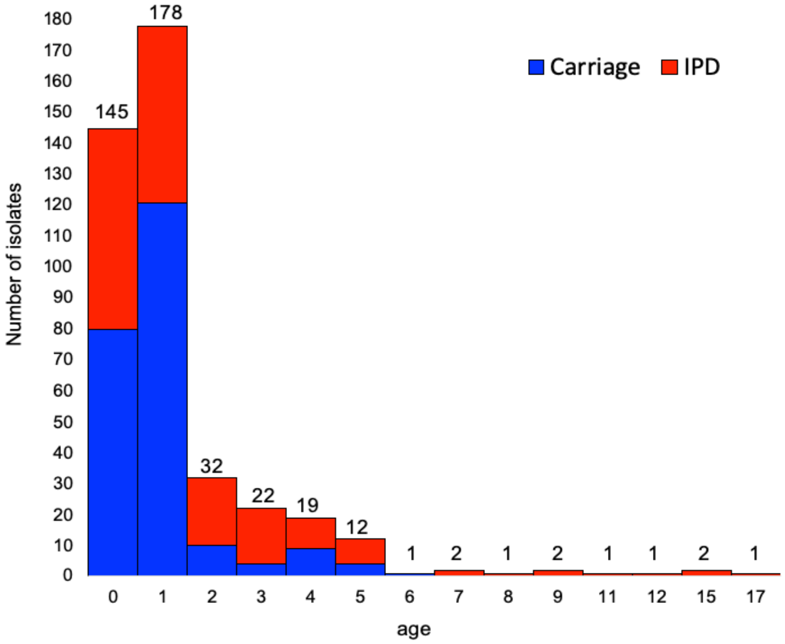
French 24F collection for WGS



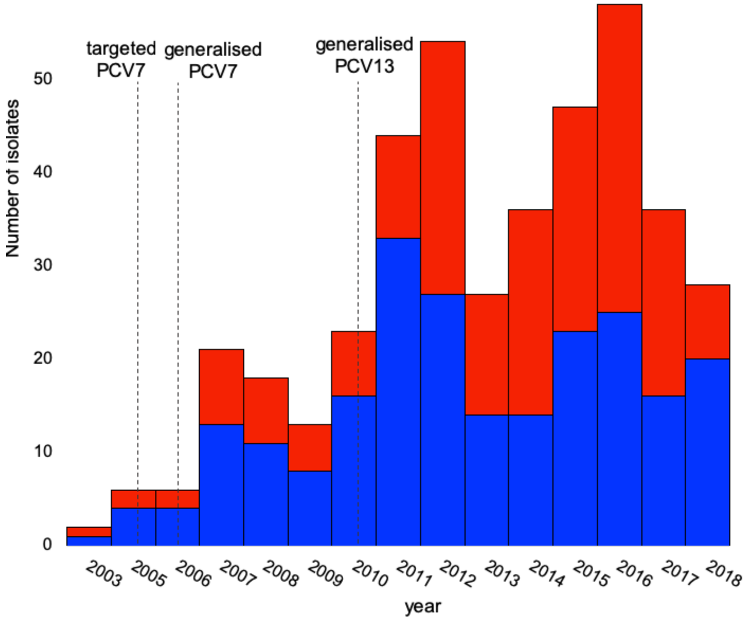
Total : n=419

Disease : n=190

Carriage: n=229



Most samples from **<2yr**

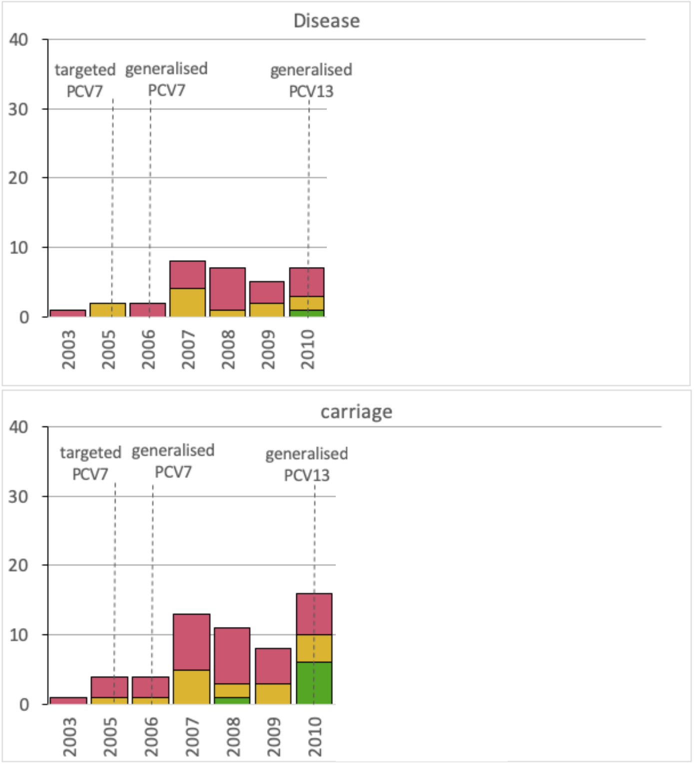


Matched carriage/disease over years

Robert Cohen and Emmanuelle Varon

What would you do with the serotype 24F pneumococcal genomes?

Clonal replacement



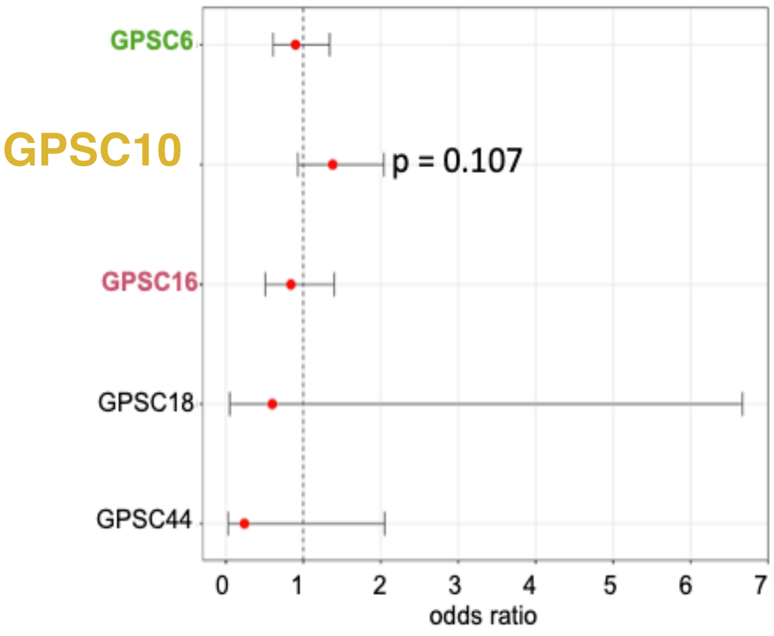
GPSC16 (S)

↓

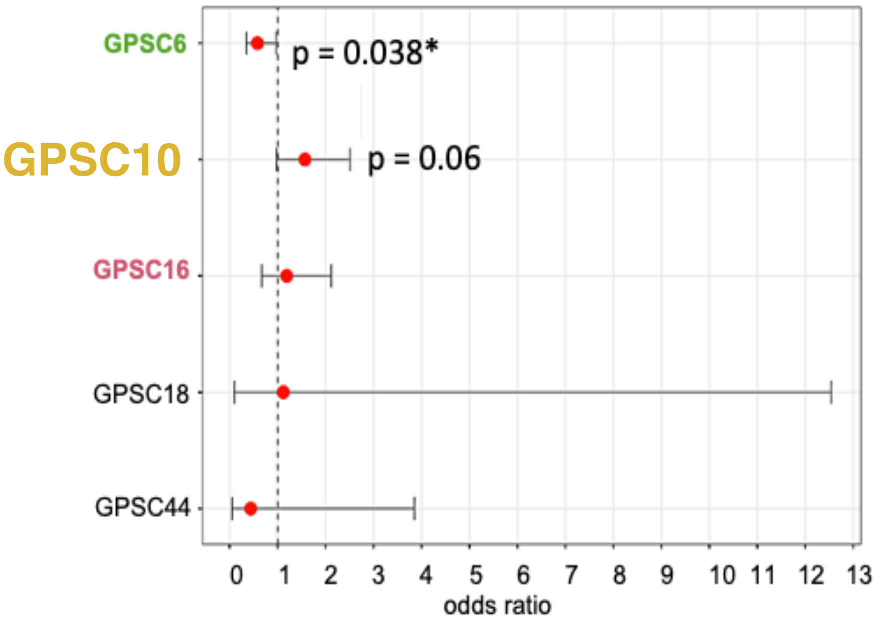
Antibiotic-selective pressure

GPSC10 - prone to cause invasive disease

Overall disease vs carriage



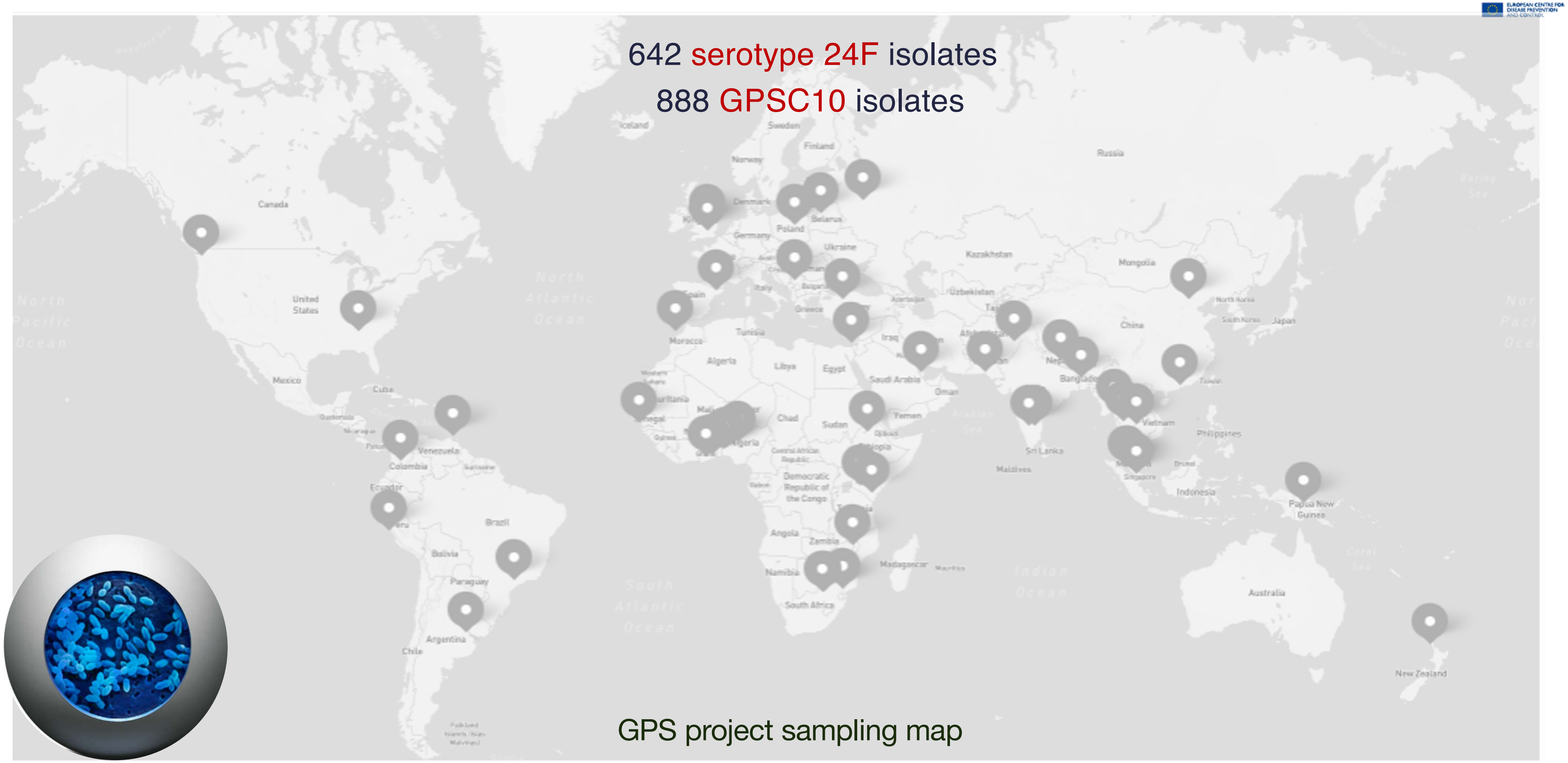
Meningitis vs Carriage



Invasive

Created in BioRender.com bio

Global context



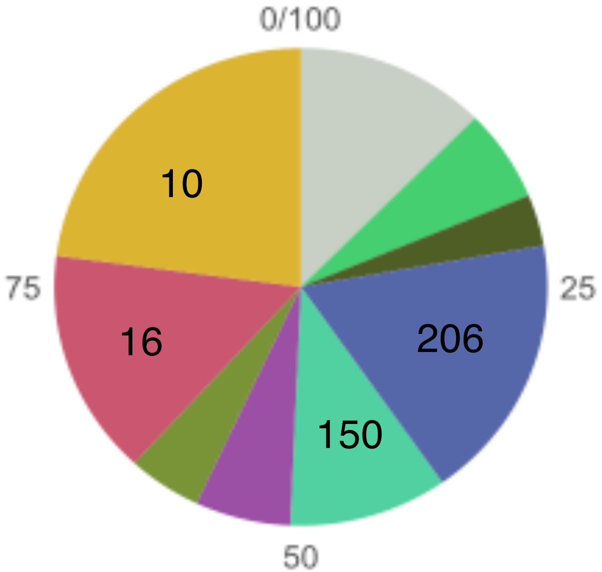
642 serotype 24F isolates

888 GPSC10 isolates

GPS project sampling map

Serotype 24F lineages

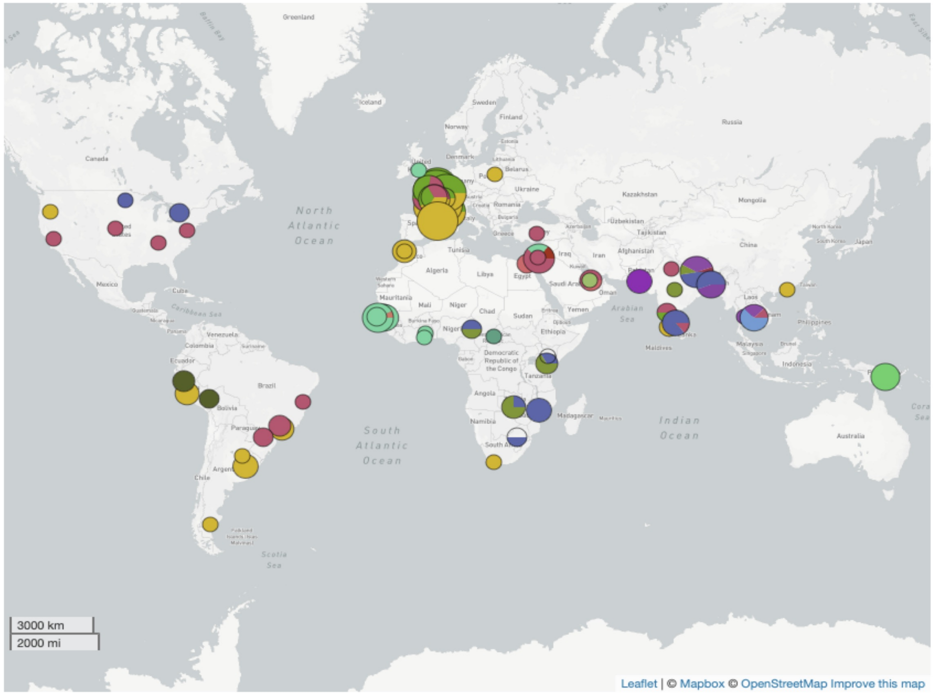
Four major lineages



30 countries

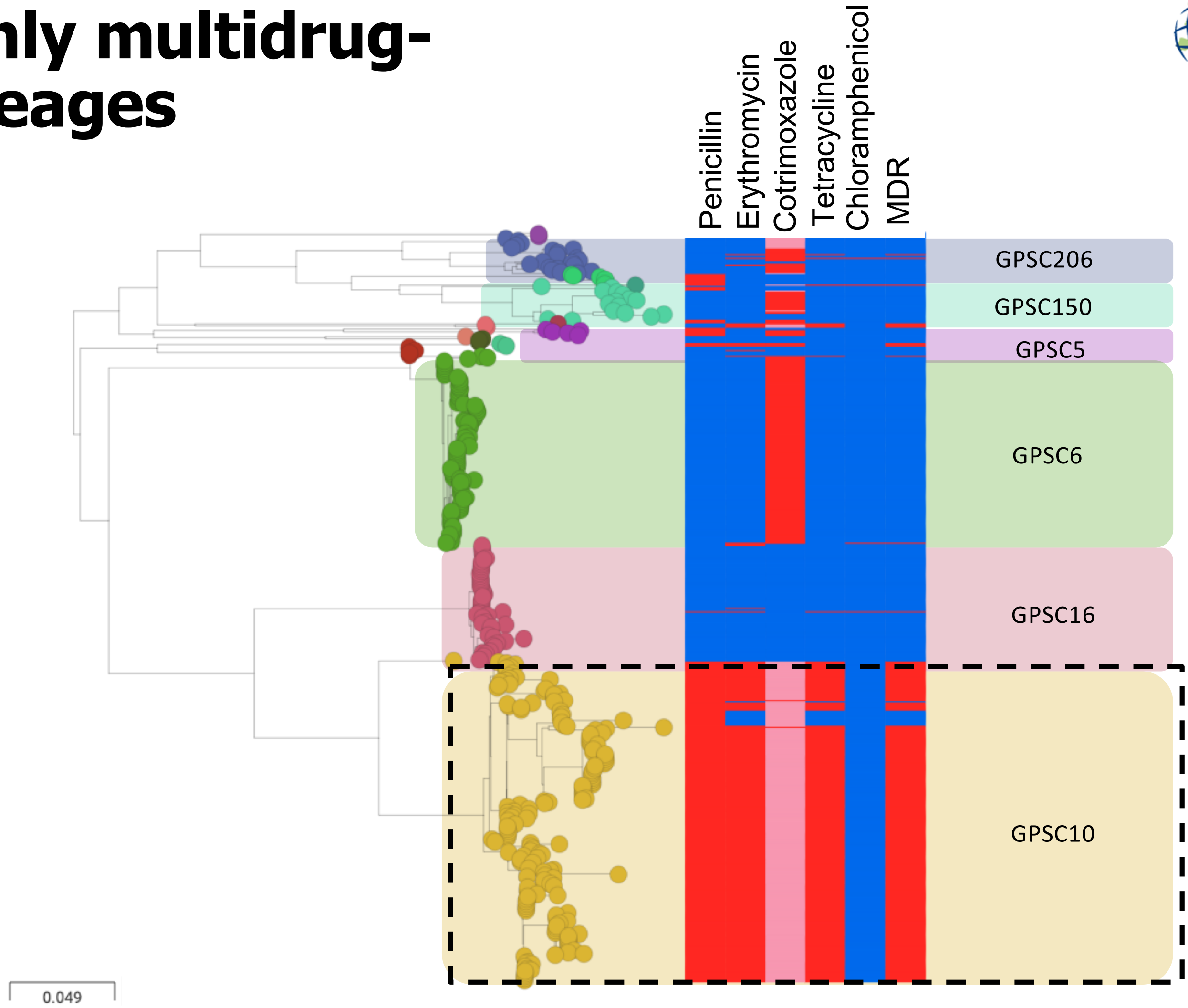
- **GPSC10** – globally spread
- **GPSC16** – globally spread
- **GPSC206** – US, Africa, Asia
- **GPSC150** – West Africa

Global spread of serotype 24F



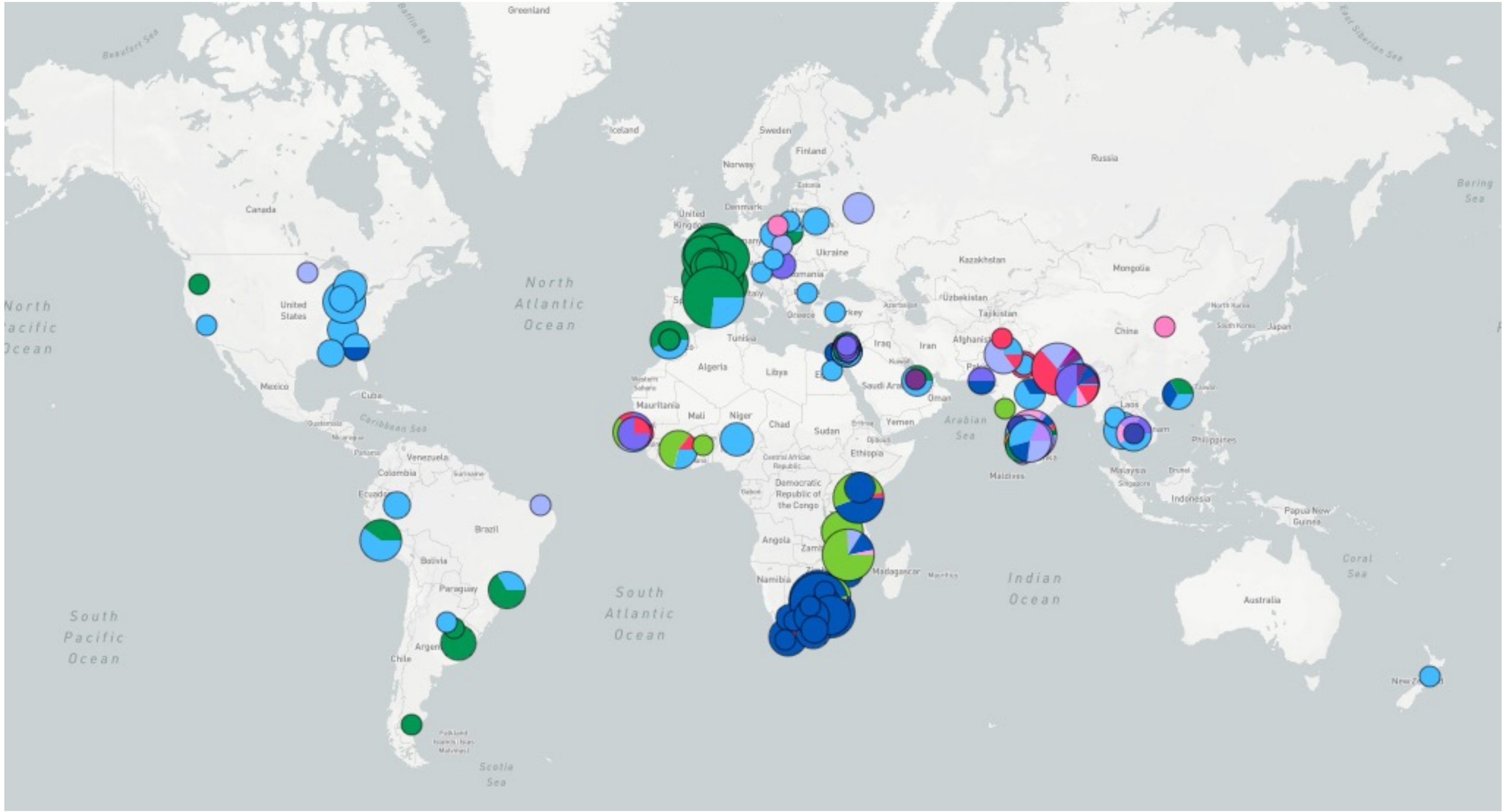
GPSC10 is the only multidrug-resistant 24F lineages

GPSC10 is resistance to penicillin, macrolide, cotrimoxazole and tetracycline.



Global GPSC10

33 countries



17 serotypes

9 PCV20 serotypes

● 3　● 6A　● 10A

● 11A　● 14　● 15B

● 19A　● 19F　● 23F

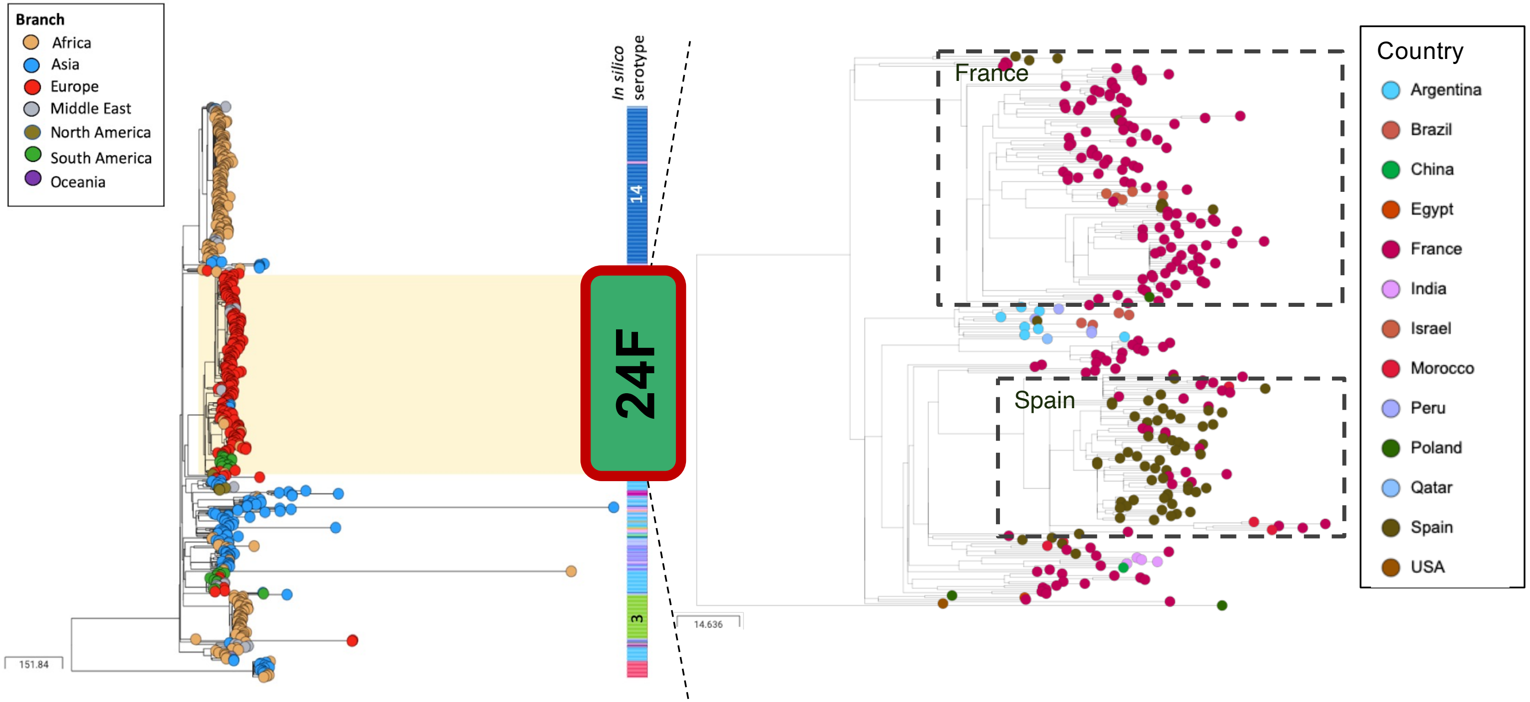
8 non-PCV20 serotypes

○ 6C　● 7B　● 13

● 15C　● 17F　● 23A

● 23B　● 24F

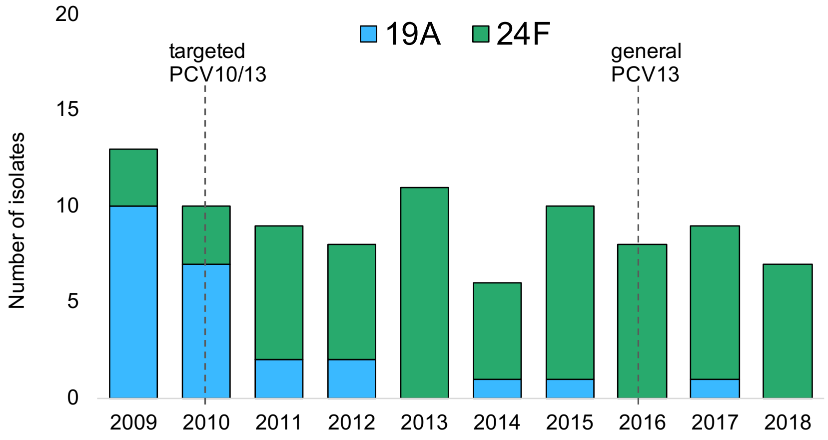
Global GPSC10 phylogeny



Lo et al 2022 The Lancet Microbe

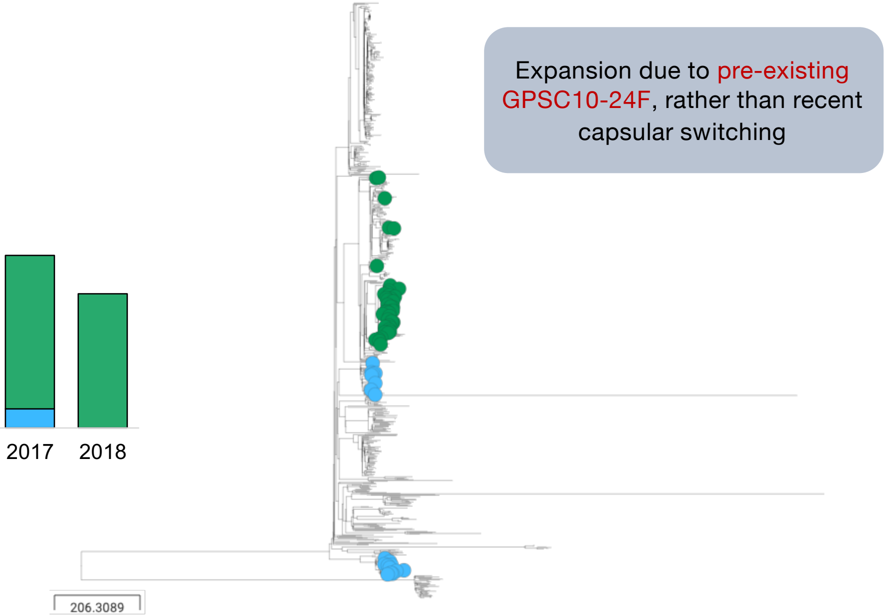
GPSC10's rapid adaptation in Spain

Rapid adaptation under vaccine-selective pressure



Carmen Muñoz Almagro

GPSC10 global phylogeny

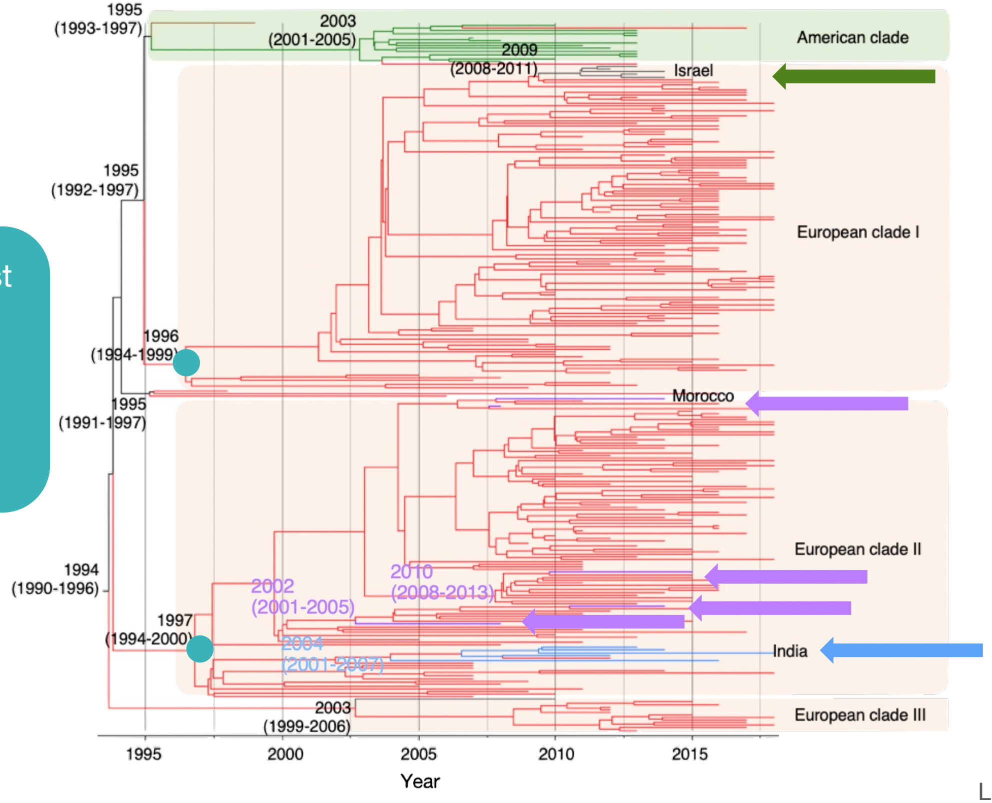


Expansion due to pre-existing GPSC10-24F, rather than recent capsular switching

Lo et al 2022 The Lancet Microbe

Origin and transmission of GPSC10-24F

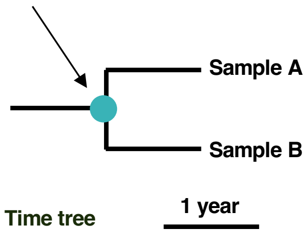
GPSC10-24F first found in three adult patients Naples Italy between 1997-1998



Lo et al 2022 The Lancet Microbe

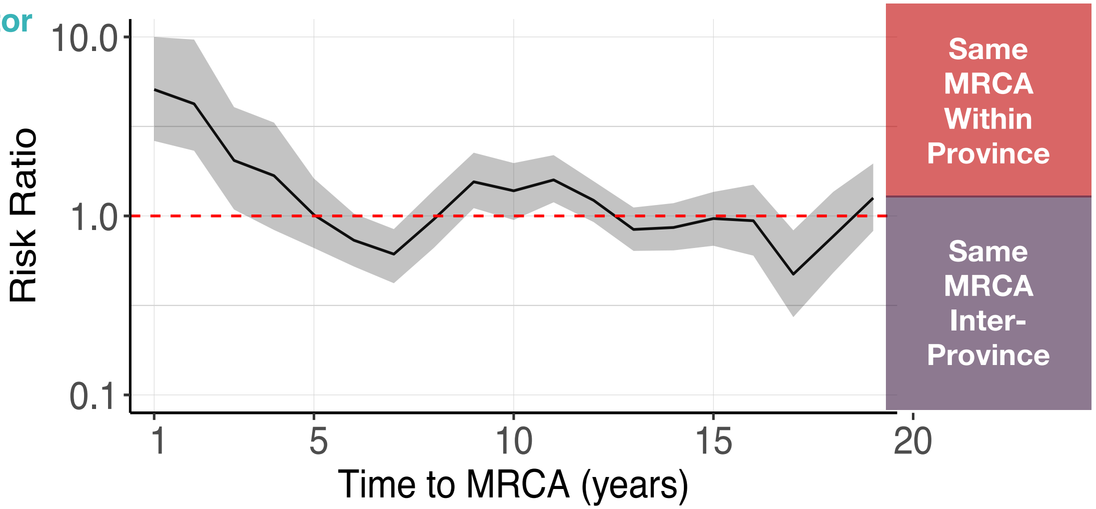
How long does it take for GPSC10-24F to spread across France?

Most recent common ancestor (MRCA)



Sophie Belman

Risk of related pairs being collected within same province in France compared to distant provinces

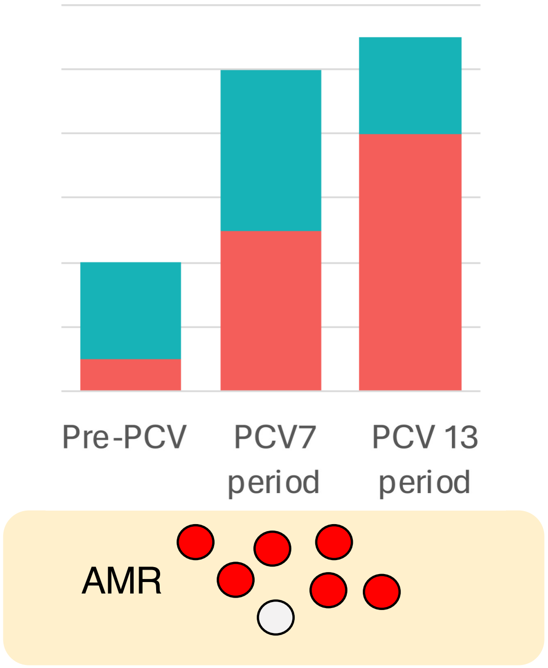


GPSC10

It is a perfect storm of high invasiveness, multidrug resistance and vaccine-evading.

It spread far and fast.

A top-5 lineage in India[1], Pakistan[2], and Nepal[3], where pneumococcal disease burden is high.



[1] Nagaraj et al 2021 Microbial Genomics
[2] Javaid et al 2024 Microbial Genomics
[3] Kandasamy et al 2022 Lancet Microbe

Gates Foundation:

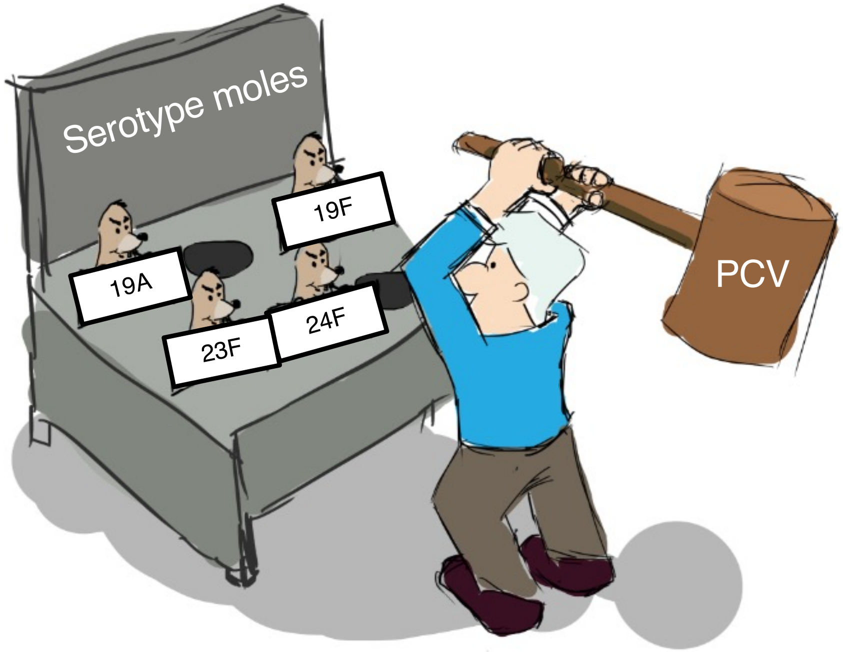
Serotype 24F is included in an upcoming vaccine formulation because of our concern about this lineage.

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

BILL & MELINDA GATES foundation

Inventprise
Inventions for global health

Next generation vaccine



Dual targets?

- Serotype
- Conserved surface proteins in super lineages like GPSC10

In summary

List of learning points in this session:

In relation to **biodiversity and population structure**

- Genomic data provides higher resolution to define pneumococcal population
- The lineage typing provides another layer of information beyond serotyping to understand vaccine evasion

In relation to **vaccine evasion**:

- Vaccine evasion is not a result of recent serotype switching, but an expansion of pneumococcal lineages that can expressed both vaccine and non-vaccine serotypes.

In summary

List of learning points in this session:

In relation to **added value** of genomics in pathogen surveillance

- Permeant digital copy of data for long-term storage
- Allow placing the local finding into global context by including genome and epidemiological data from elsewhere
- Spatiotemporal analysis allow us to estimate the emergence and transmission of a particular lineage
- Inform vaccine design

Further reading

Specific further reading for this session (e.g. Book, article, blog, video, etc.)

Vaccine failure and breakthrough

Feroze et al 2023 J Bio Chem Novel pneumococcal capsule type 33E results from the inactivation of glycosyltransferase WciE in vaccine type 33F (Upcoming PCV including 33F is predicted not to be protected against 33E)

Corcoran et al 2021 Vaccine Using genomics to examine the persistence of Streptococcus pneumoniae serotype 19A in Ireland and the emergence of a sub-clade associated with vaccine failures (genetic determinants might explain vaccine failure)

Further reading

Specific further reading for this session (e.g. Book, article, blog, video, etc.)

Serotype replacement

Lo et al 2019 Lancet Infect Dis Emergence of a multidrug-resistant and virulent Streptococcus pneumoniae lineage mediates serotype replacement after PCV13: an international whole-genome sequencing study

Lekhuleni et al 2024 Nat Comm Impact of pneumococcal conjugate vaccines on invasive pneumococcal disease-causing lineages among South African children

Further reading

Specific further reading for this session (e.g. Book, article, blog, video, etc.)

GPS project resources

21k pneumococcal genome database (tutorial video)
GPS visualisation: Interactive display of serotypes/lineages across countries after introduction of PCV13 (tutorial video)
107 serotypes
PCV formulations
GPS bioinformatics training
GPS pipeline Nat Comm paper

References

Lees et al 2019 Genome Res PopPUNK https://pubmed.ncbi.nlm.nih.gov/30679308/

Gladstone et al 2019 Ebiomedicine Pneumococcal population structure
https://pubmed.ncbi.nlm.nih.gov/31003929/

Lo et al 2022 Lancet Microbe GPSC10 mediates serotype replacement
https://pubmed.ncbi.nlm.nih.gov/35985351/

Hung et al 2024 BioRxiv GPS pipeline
https://www.biorxiv.org/content/10.1101/2024.11.27.625679v1

King et al 2024 GPSC vs MLST vs cgMLST vs LIN code
https://pmc.ncbi.nlm.nih.gov/articles/PMC11353345/

Ouldali et al 2018 The Lancet Infect Dis – serotype replacement in France mediated by 24F
https://pubmed.ncbi.nlm.nih.gov/30049623/

References

Wahl et al 2018 Global Pneumococcal disease burden
https://pubmed.ncbi.nlm.nih.gov/29903376/

Nagaraj et al 2021 Baseline population structure of pneumococci in India before PCV introduction
https://pubmed.ncbi.nlm.nih.gov/34494953/

Javaid et al 2024 Pneumococcal lineages circulating in Pakistan and early impact of PCV
https://pubmed.ncbi.nlm.nih.gov/38270581/

Kandasamy et al 2022 Lancet Microbe PCV impact on pneumococcal populations in children in Nepal
https://pubmed.ncbi.nlm.nih.gov/35779566/

Acknowledgements

The creation of this training material was commissioned by ECDC to Institut Pasteur with the direct involvement of Dr Stephanie Lo
slo@ebi.ac.uk