



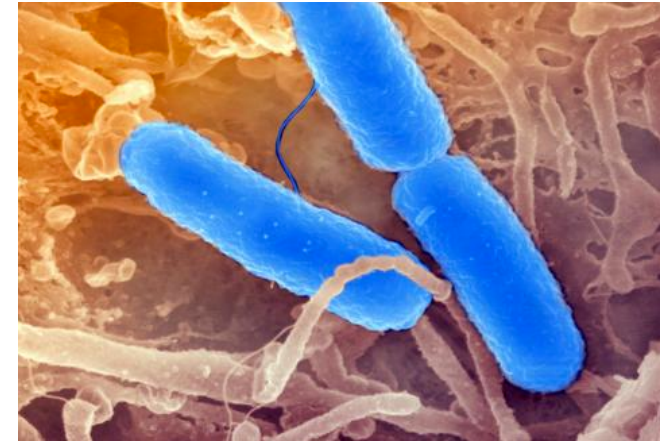
# Whole-genome analysis of the epidemic dysentery bacillus (*Shigella dysenteriae* type 1)

François-Xavier WEILL, MD, PhD  
Enteric Bacterial Pathogens Unit  
Infection & Epidemiology Department



# Shigella

- Bacillary dysentery (bloody diarrhea)
- Human reservoir
- 100 million cases, 100,000 death/year
- Antimicrobial resistance (first reports 1950s)
- Four serogroups (*S. dysenteriae*, *flexneri*, *boydii* and *sonnei*), > 50 serotypes



## *Shigella dysenteriae* type 1

- Severe infections (Shigatoxin)
- Large outbreaks in developing countries
- Wars or civil conflicts, climatic events, famine, refugees

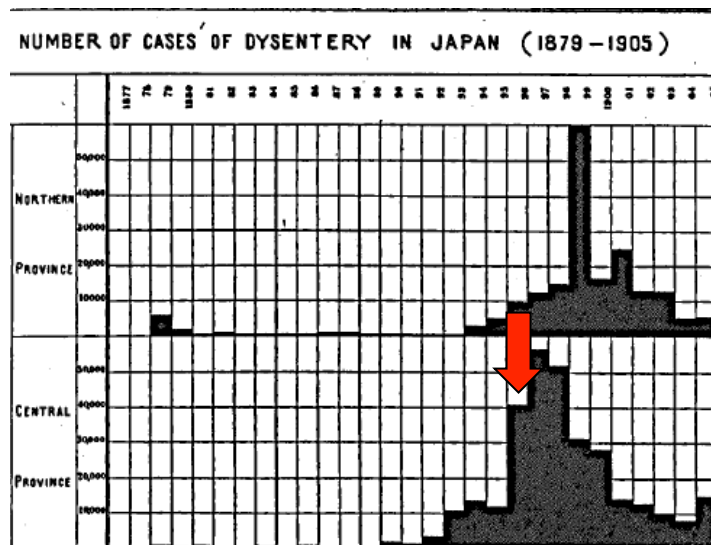


# History

## Before the advent of microbiology

- France 1738, 1742, 1778-1779, 1792, 1857
- Sweden 1772-1773, 1808-1809, 1857
- Ireland 1849-1850
- Crimean war 1853-1856
- French Prussian war 1870

## 1897 Kyoshi Shiga, Japan



Jun-Dec 1897  
90,000 cases  
20,000 deaths



Kiyoshi Shiga, Ueber den Erreger der Dysenterie in Japan. 599

*Nachdruck verboten.*

**Ueber den Erreger der Dysenterie in Japan.**  
 [Aus dem Institute für Infektionskrankheiten des Herrn Prof.  
 Dr. Kitasato zu Tokio.]  
 Vorläufige Mitteilung.  
 Von  
**Dr. Kiyoshi Shiga,**  
 Assistenten am Institute.

Die Dysenterie herrscht in Japan seit 30 Jahren alljährlich fast in allen Provinzen. Auch in diesem Jahre sind vom Juni bis Dezember beinahe 90000 Personen erkrankt mit über 20000 Todesfällen.

Tokio, den 10. Dezember 1897.

# History II

## Pre WWI

- Germany, Russia, France, England, US : small outbreaks in garrisons, plants and asylums

## WWI

- Gallipoli or Dardanelles campaign, Apr 1915- Jan 1916  
Most of the 120,000 casualties
- Western front  
clusters but more often *S. flexneri*

## WWII

- Eastern Europe and Middle East

## Post WWII

- Bangladesh Liberation war 1971-1972
- Central America, 1968-1972 (500,000 cases and 20,000 deaths)
- Central Africa 1979-1980s (100,000 cases and 5-10,000 deaths)
- Africa 1991-2006 (Rwanda, RDC, Sierra Leone, CAR, Zambia, Liberia, ...)

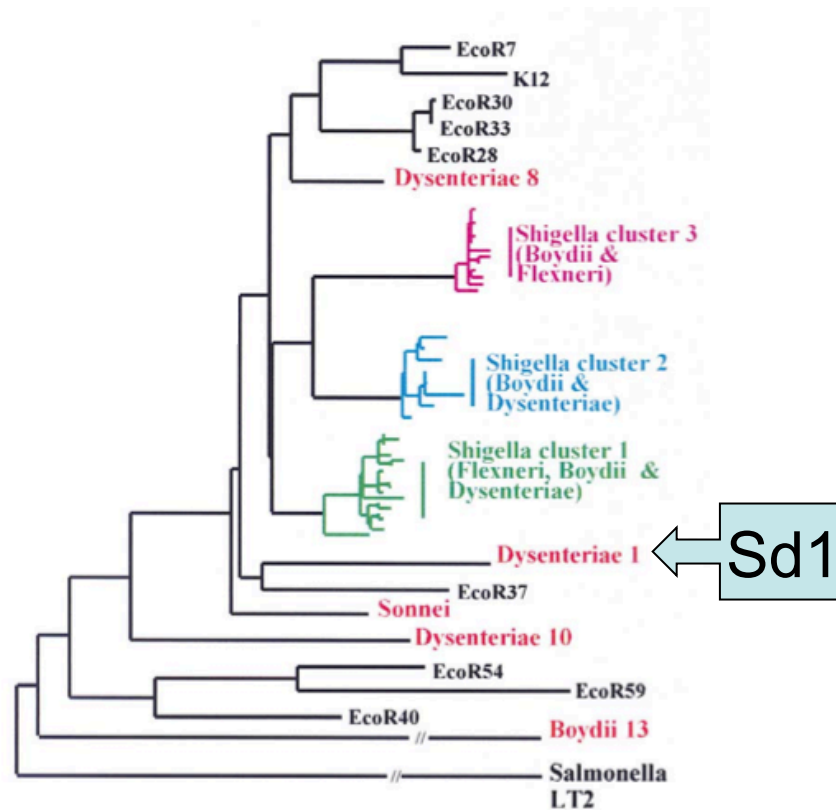


Royal Mail, May 2015

Multidrug  
resistance

# Taxonomy

« *Shigella* » are *Escherichia coli* lineages



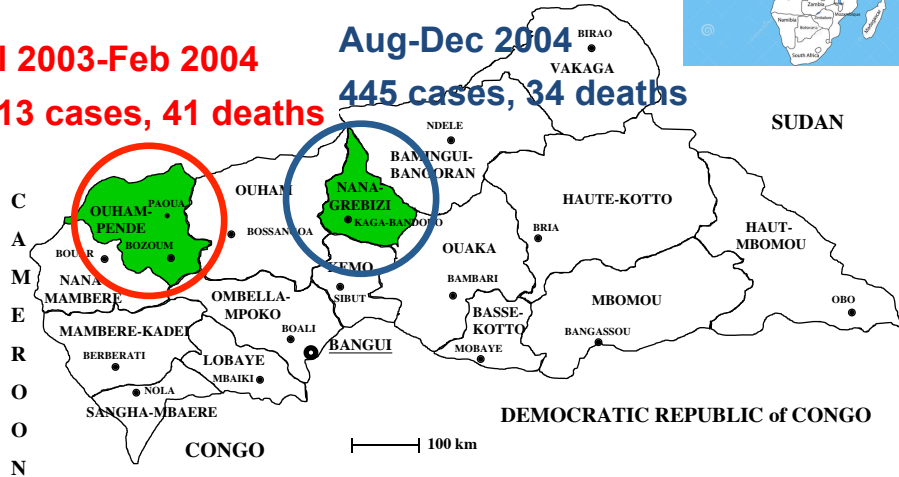
Lan & Reeves, Microb Infect 2002

Multilocus Sequence Typing:

# Why Sd1 ?

Jul 2003-Feb 2004  
2013 cases, 41 deaths

Aug-Dec 2004  
445 cases, 34 deaths

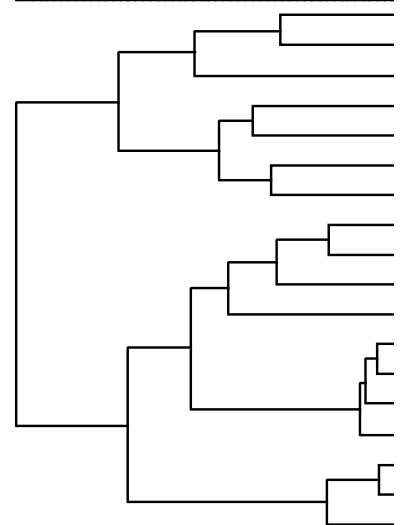


## Molecular epidemiology of multidrug-resistant *Shigella dysenteriae* type 1 causing dysentery outbreaks in Central African Republic, 2003–2004

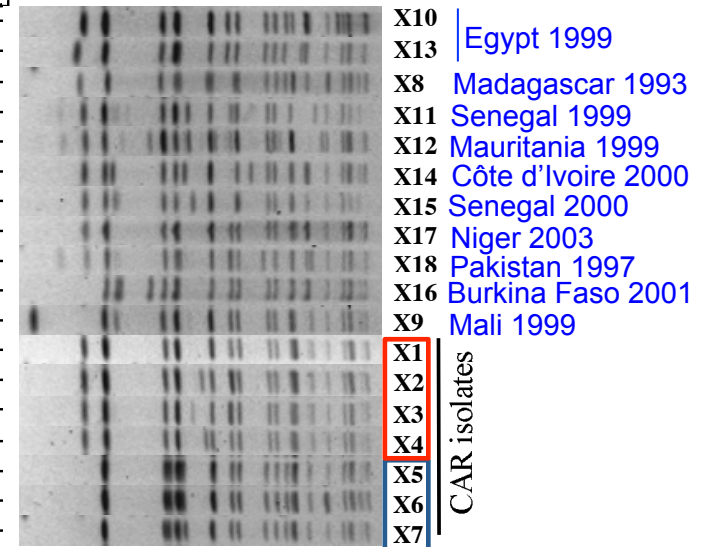
Raymond Bercion<sup>a</sup>, Marie Demartin<sup>b</sup>, Carlos Recio<sup>c</sup>, Peguy-Martial Massamba<sup>c</sup>, Thierry Frank<sup>a</sup>, Josep M. Escribà<sup>d</sup>, Francine Grimont<sup>b</sup>, Patrick A.D. Grimont<sup>b</sup>, François-Xavier Weill<sup>b,\*</sup>

**Conclusion:**  
Two different strains according to PFGE and plasmid profiling

Dice (Tol 0.5%-0.5%) (H>0.0% S>0.0%) [0.0%-100.0%]  
Relative similarity (%)  
50 60 70 80 90 100



PFGE-XbaI



**Whole-genome sequencing**

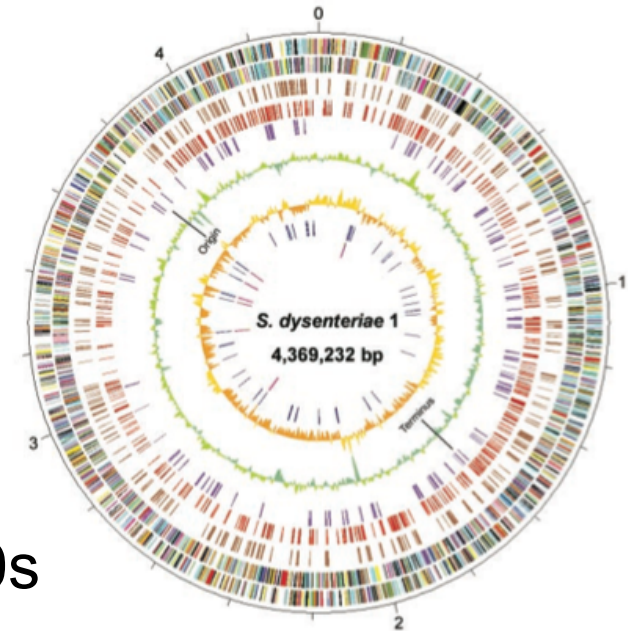
# Sequenced Sd1 genomes

- **Sd197** (Yang et al., Nucleic Acids Res 2005)
  - 4.37 Mb
  - 4557 CDSs
  - 285 pseudogenes
  - 623 intact insertion sequences (12%)
  - Virulence plasmid 182 kb

China, 1950s

- **Sd1617** (Vongsawan, FEMS Microbiol. Lett 2015)
  - 4.3 Mb
  - 6,505 ORFs
  - 1,790 insertion sequences
  - Virulence plasmid 177 kb

Guatemala, 1968-1969



## WGS study

- Number of genomes: 56
- Period of time: 1950-2003
- Country coverage : 9 (Bangladesh, 30 isolates !)
- SNPs detected : 989

### RESEARCH ARTICLE

### Open Access

## Genomic analysis of the emergence of 20th century epidemic dysentery

Laurence Rohmer<sup>1\*</sup>, Michael A Jacobs<sup>1</sup>, Mitchell J Brittnacher<sup>1</sup>, Christine Fong<sup>1</sup>, Hillary S Hayden<sup>1</sup>, Didier Hocquet<sup>2,3</sup>, Eli J Weiss<sup>1</sup>, Matthew Radey<sup>1</sup>, Yves Germani<sup>4</sup>, Kaiser Ali Talukder<sup>5</sup>, Anthony J Hager<sup>1</sup>, John M Kemner<sup>1</sup>, Elizabeth H Sims-Day<sup>1</sup>, Susana Matamouros<sup>1</sup>, Kyle R Hager<sup>1</sup> and Samuel I Miller<sup>1,6,7</sup>

### Results

- Most common recent ancestor dated 1924
- Lack of consistency between phylogeny and geography
- The Central American outbreak strain imported from Africa in 1960

### Conclusions

Sd1 disseminated after WWI and became ubiquitous due to long term carriage

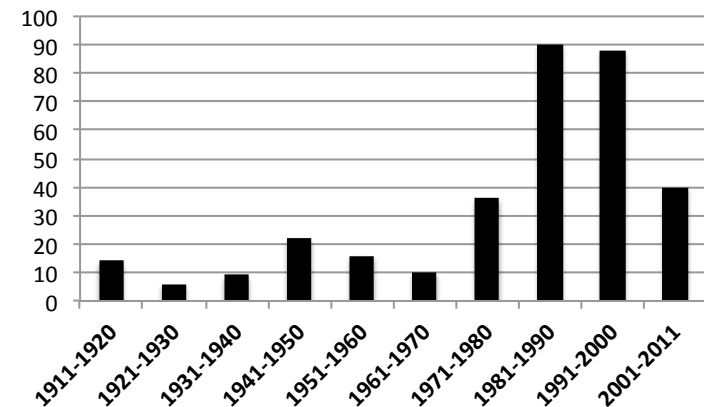
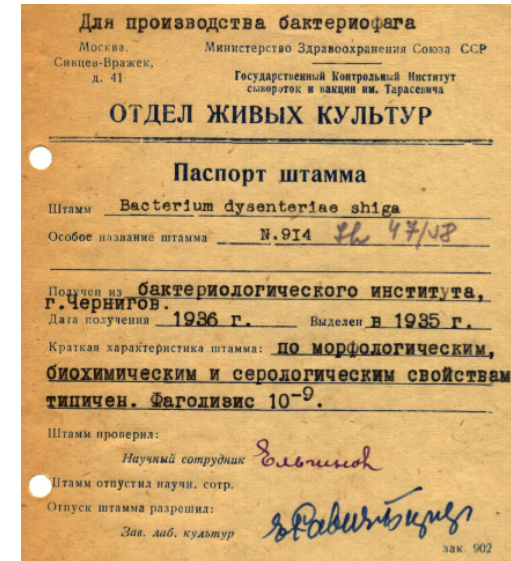
**Not consistent with isolation by Shiga in 1897 and historical data**



# Objectives of our study

- Population structure & transmission patterns
- Spatio-temporal analysis
- Linkage between the different outbreaks
- Emergence of antimicrobial resistance

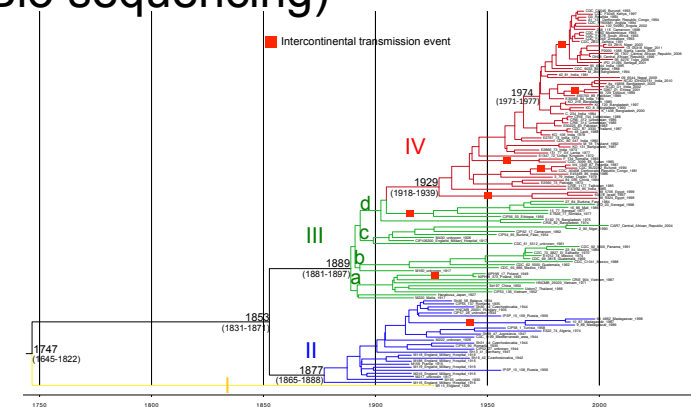
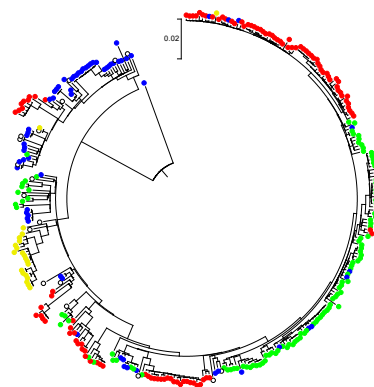
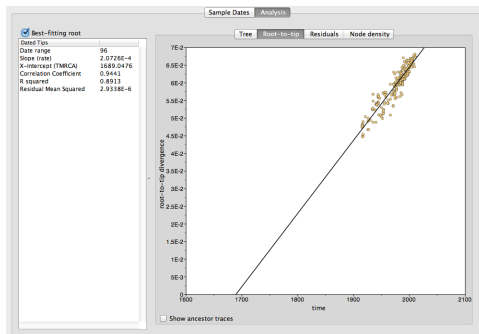
Exceptional collection of 325 isolates  
plus 6 published genomes,  
66 countries, 1915-2011



1915-1960,  $n=67$

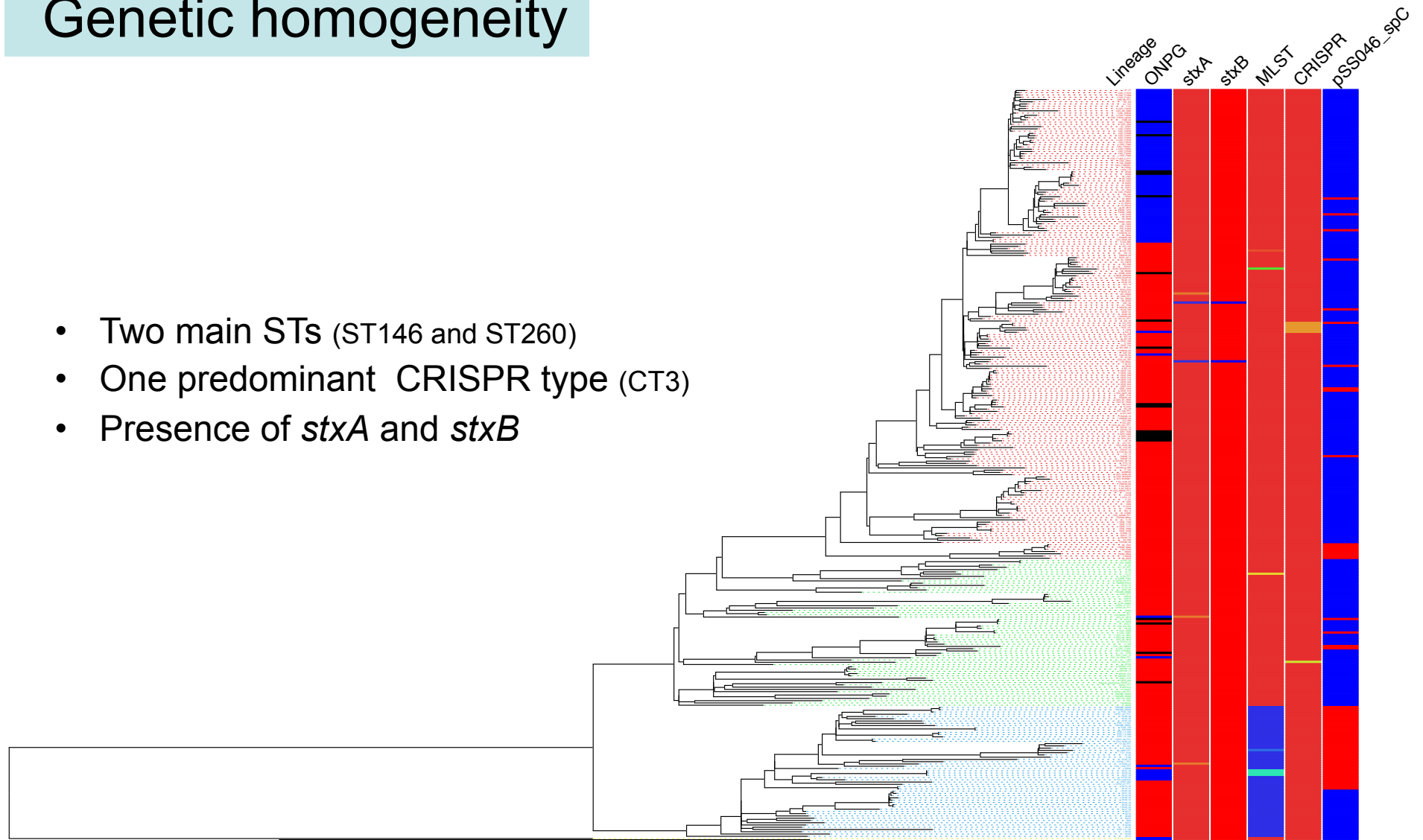
# Methodology

1. Illumina sequencing (100-150 bp PE short reads)
2. Mapping against reference genome Sd197 (or Sd1617)
3. SNPs calling and filtering (SMALT), assembling (Velvet, CLC)
4. Phylogenetic and phylogeographic analyses and estimation of the divergence times of the different lineages/strains (RAxML after Gubbins, Path-O-Gen, Beast)
5. Extraction of antibiotic resistance genes and their genetic support (Resfinder on read assemblies, 454 sequencing, PacBio sequencing)



# Genetic homogeneity

- Two main STs (ST146 and ST260)
- One predominant CRISPR type (CT3)
- Presence of *stxA* and *stxB*



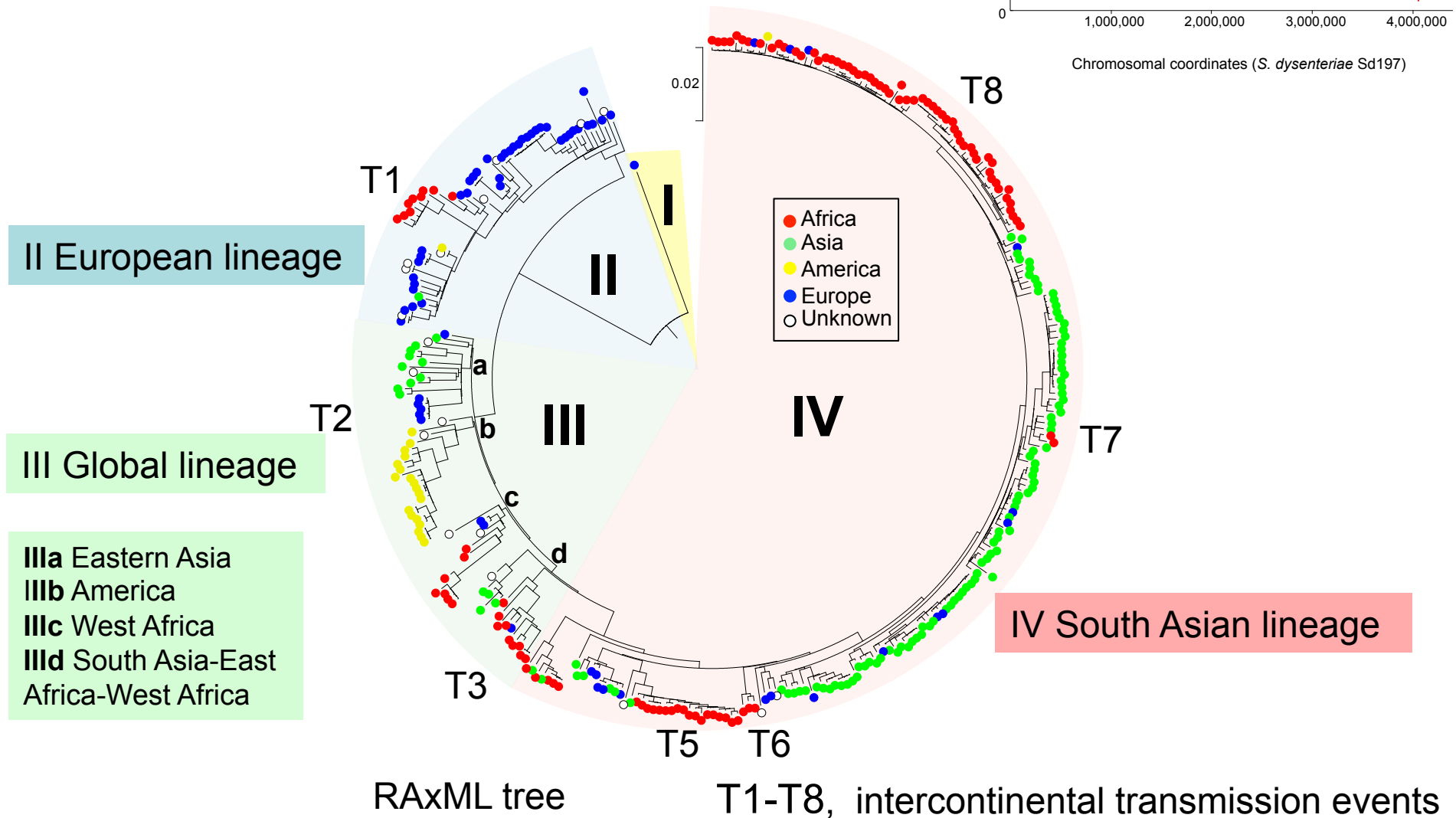
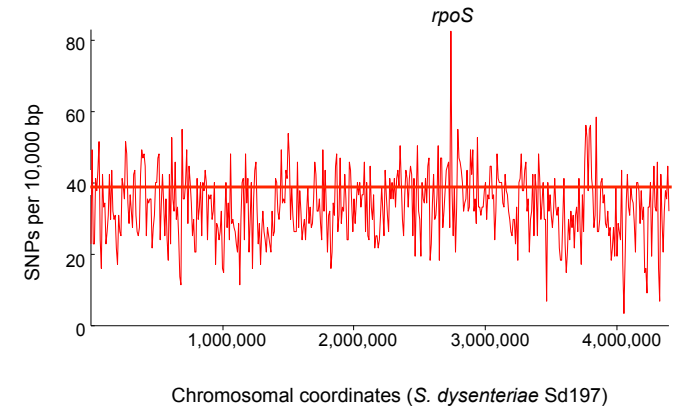
## Keys

(ONPG) Positive Negative Unknown (MLST) ST146 ST5160 ST5159 ST5157 ST260 ST5158 ST5161

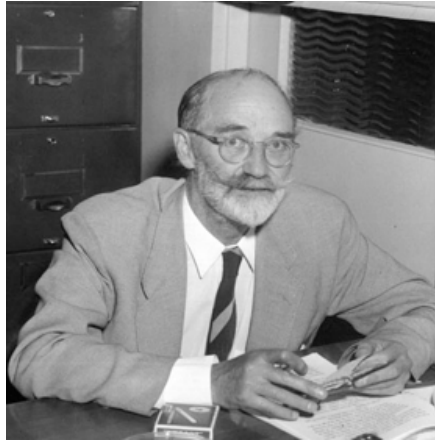
(*stxA*, *stxB*, pSS046\_spC) Presence Absence SNP (CRISPR) 3 3var1 32

# Phylogeny

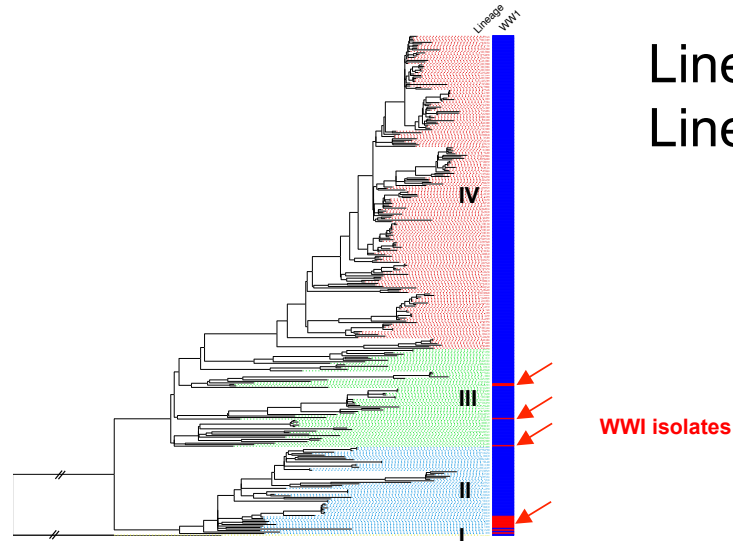
14,677 SNPs distributed on 85.6% of the Sd197 chromosome (39 SNPs every 10,000 nt)



# 14 WWI isolates



Capt. E.G.D. Murray



Lineage II,  $n=10$

Lineage III,  $n=4$

IIIa,  $n=1$

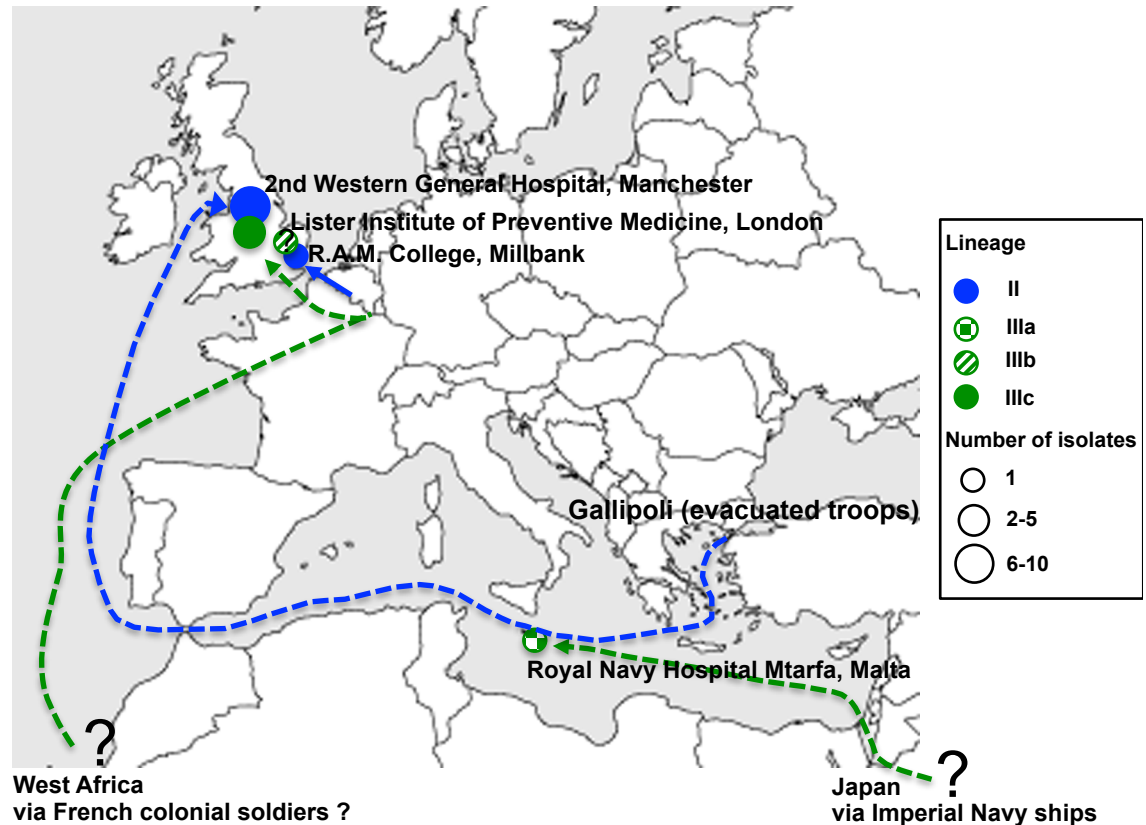
IIIb,  $n=1$

IIIc,  $n=2$

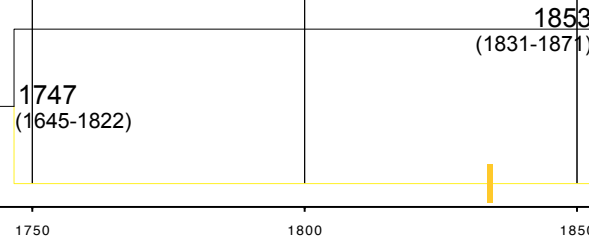
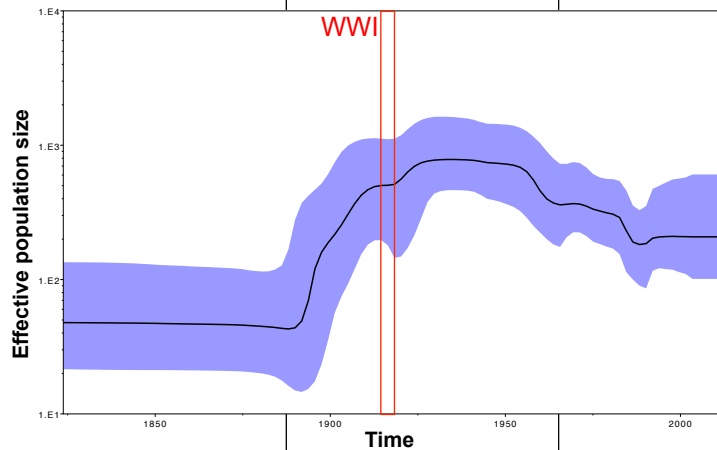
II : Gallipoli

IIIc : African colonial soldiers ?

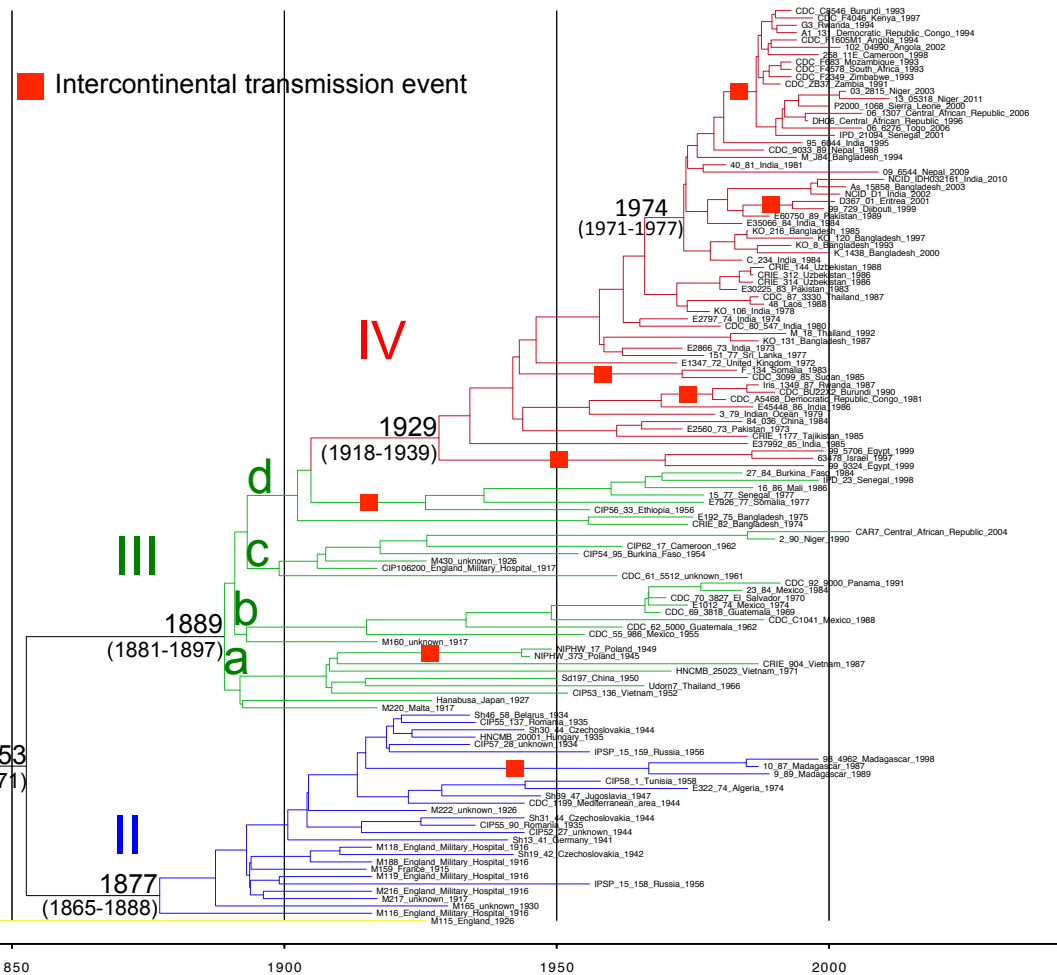
IIIa : Japanese sailors ?



# BEAST analysis



■ Intercontinental transmission event



125 isolates; 10,798 SNPs

Log normal relaxed clock rate

Bayesian skyline population size

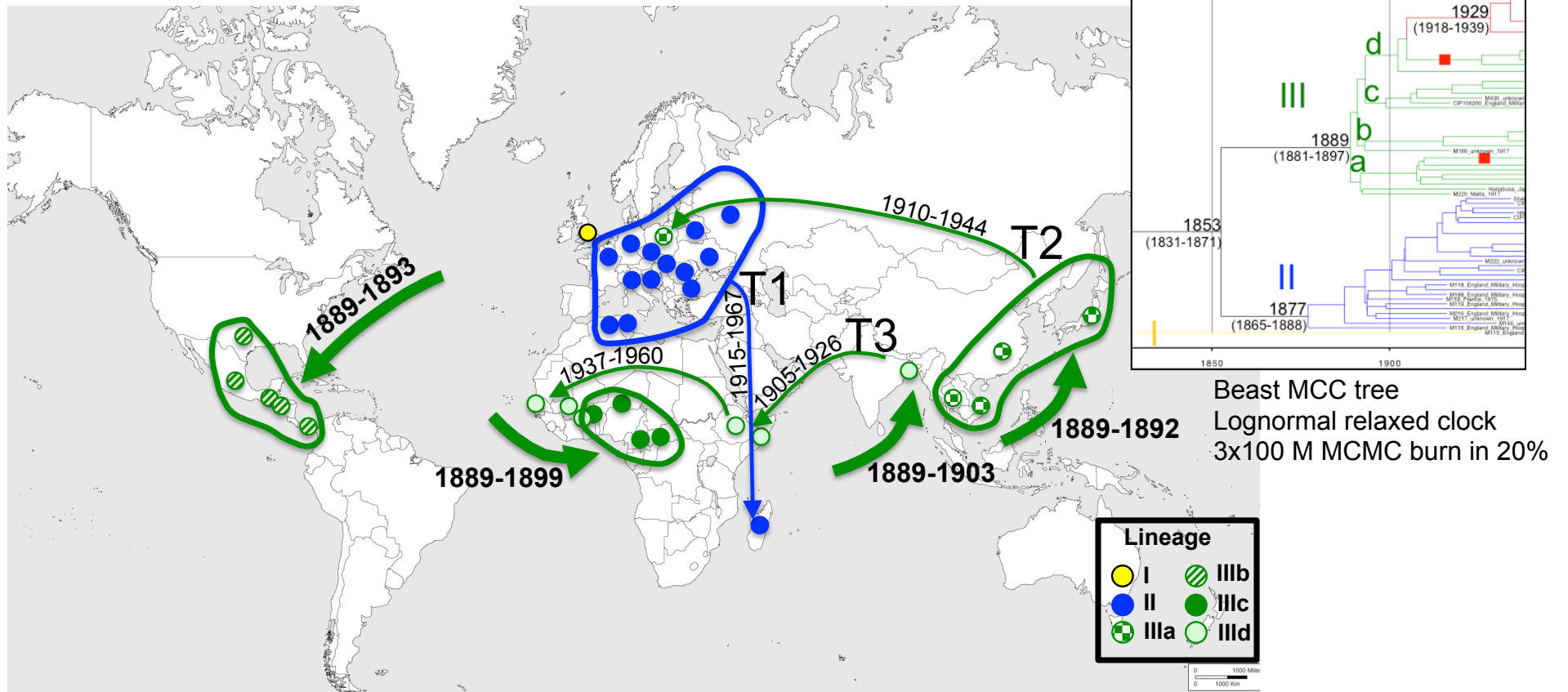
GTR+Gamma

MCMC 100 million x3

$8.7 \times 10^{-7}$  substitutions site<sup>-1</sup> year<sup>-1</sup>  
 [95% HPD =  $7.6 \times 10^{-7} - 9.9 \times 10^{-7}$ ]

Divergence times confirmed on 10 other subsamples of 125 isolates

# Transmission patterns of lineages II and III

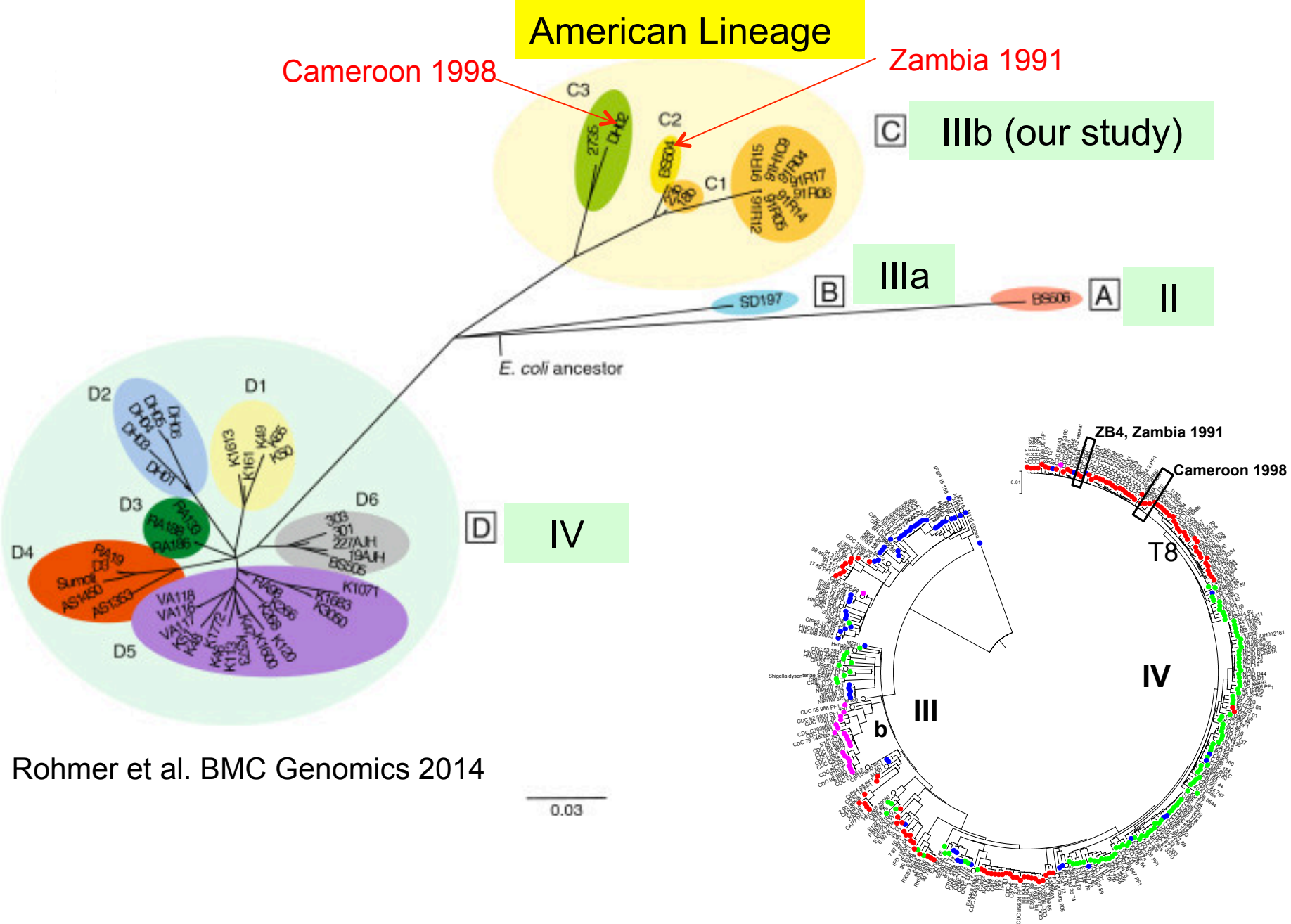


Lineage II: Europe and North Africa, extincted 1950-1970 (except in Madagascar)

Lineage III: Global distribution in less than 2 decades (1889-1903)

European emigration, colonization, trading routes facilitated by the opening of the Suez Canal and the development of steamships

American lineage imported from Africa in 1960 (Rohmer et al.) ???



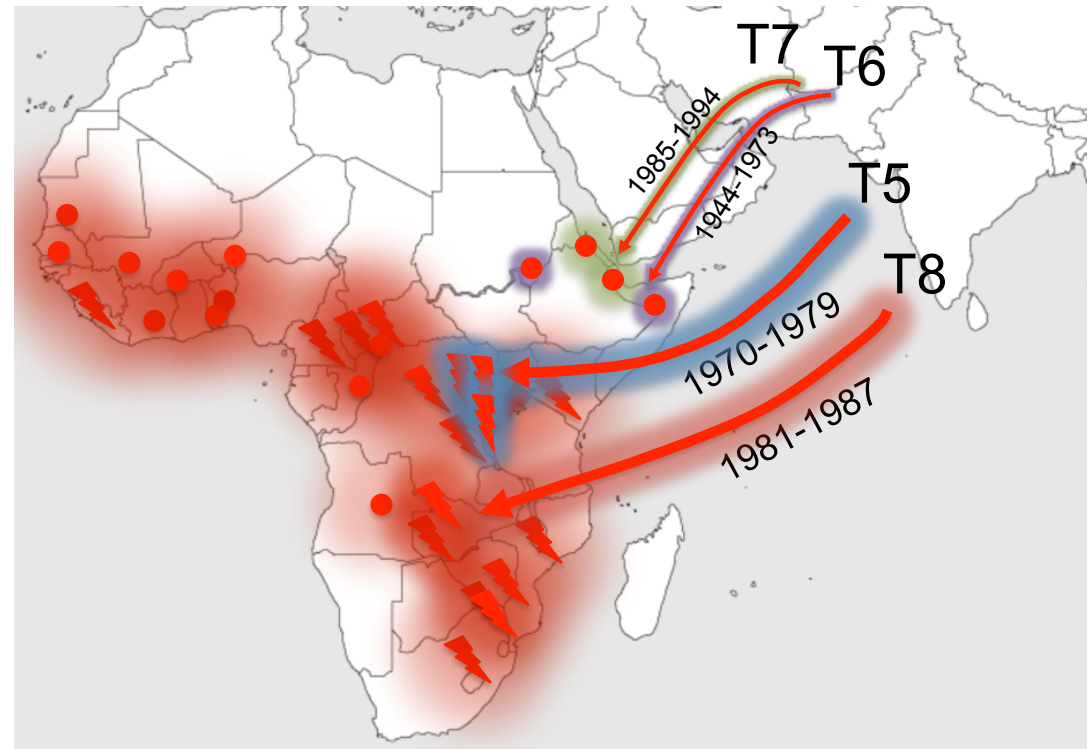
Rohmer et al. BMC Genomics 2014





## Transmission patterns of lineage IV

- Emerged in 1929 (95% CI 1918-1939)
- Indian subcontinent with transmission waves to :
  - East or Central Africa
  - South East Asia
  - Central Asia

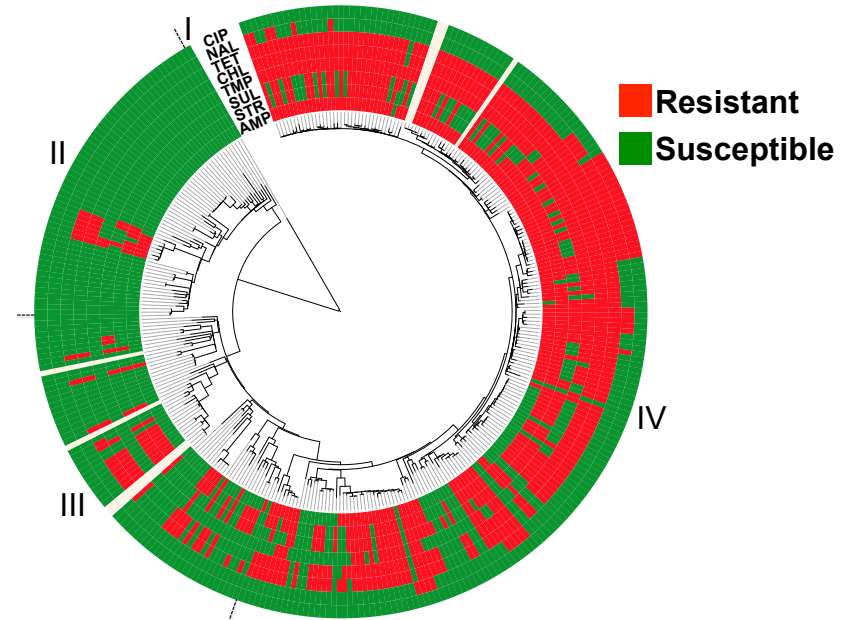


- Sequence divergence times correlated with published data:

e.g., T5, the « Zairian outbreak » 100,000 cases was first reported at the end of 1979 in northeastern RDC

- T5 was geographically and temporally restricted (Great Lakes until 1990)
- T8 disseminated across Africa during the 1990s

# Antibiotic resistance I



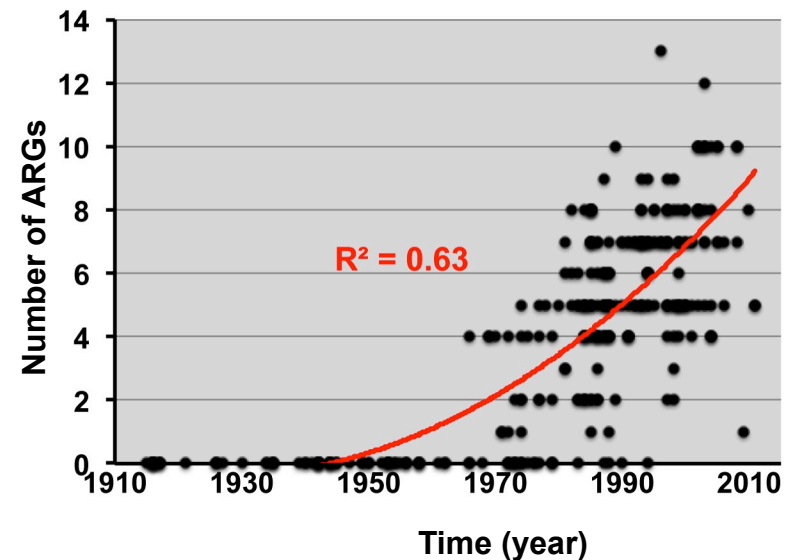
- Key element in Sd1 evolution
- Started during the 1960s independently in Asia, Africa and America
- Rapid and massive accumulation of antibiotic resistance genes (ARGs)

1915-1960: 100% (67/67) susceptible isolates

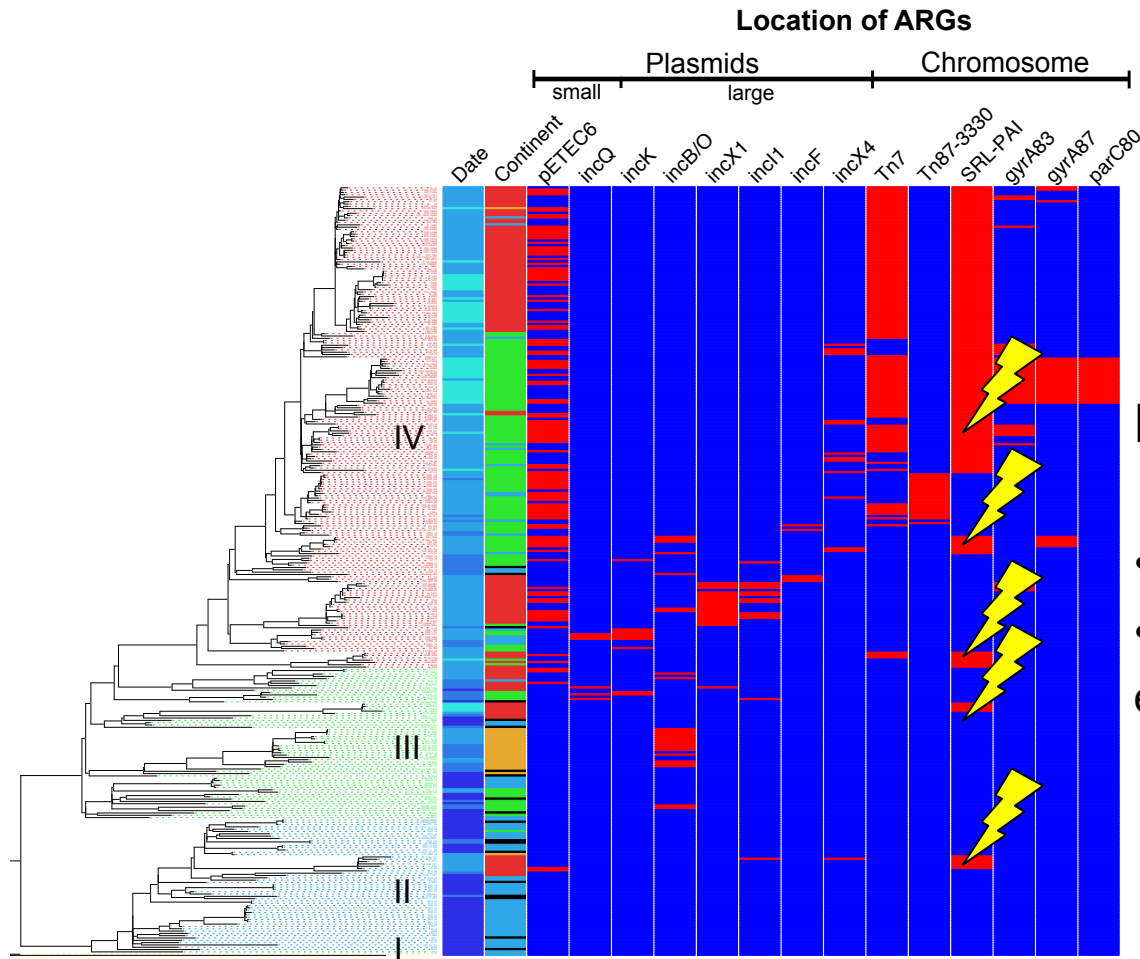
1991-2011: <1% (1/123)

- Following clinical practice

From streptomycin, sulfonamides, tetracycline, chloramphenicol, ampicillin, cotrimoxazole, nalidixic acid to ciprofloxacin



# Antibiotic resistance II



**Keys**  
 (Continent) Africa America Asia Europe unknown  
 (Date) 1915-1959, 1960-1979, 1980-1999, 2000-2011  
 (Others) Presence Absence

## Phase 1 (1960s-1992)

### Multidrug resistant plasmids

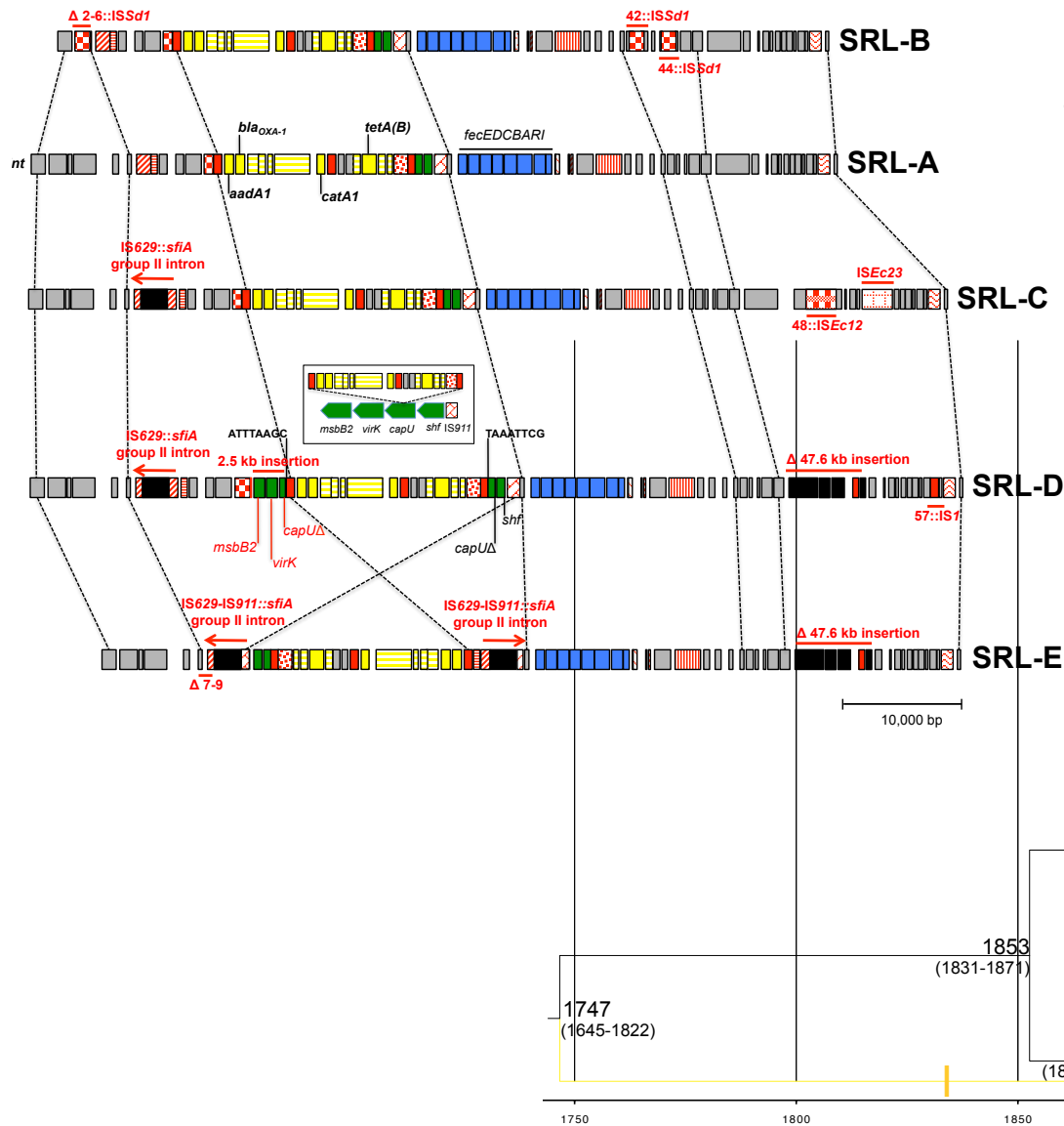
- Small (<10 kb) to large size (130 kb)
  - Local acquisition
- e.g.,
- IncB/O in Central America
  - IncX1 then I1 in Central Africa
  - IncK in Asia

## Phase 2 (after mid 1970s)

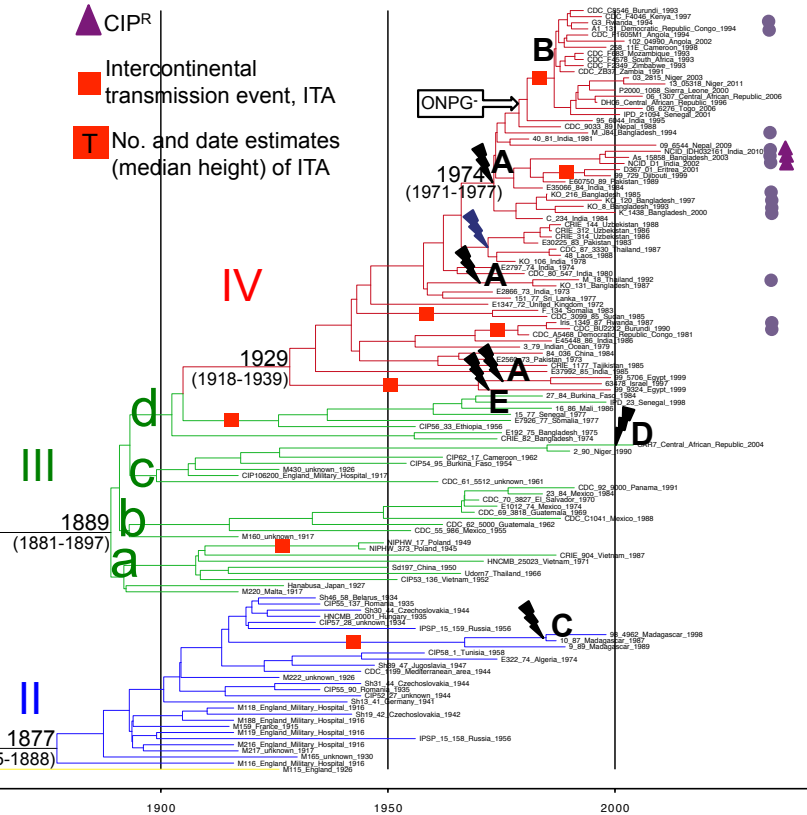
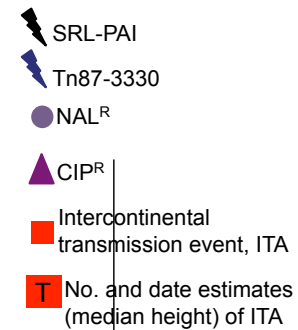
### Chromosomal determinants

- Transposons: Tn7 and Tn87-3330
- **Genomic island: SRL-PAI** ⚡  
acquired 6 times
- *gyrA* and *parC* mutations

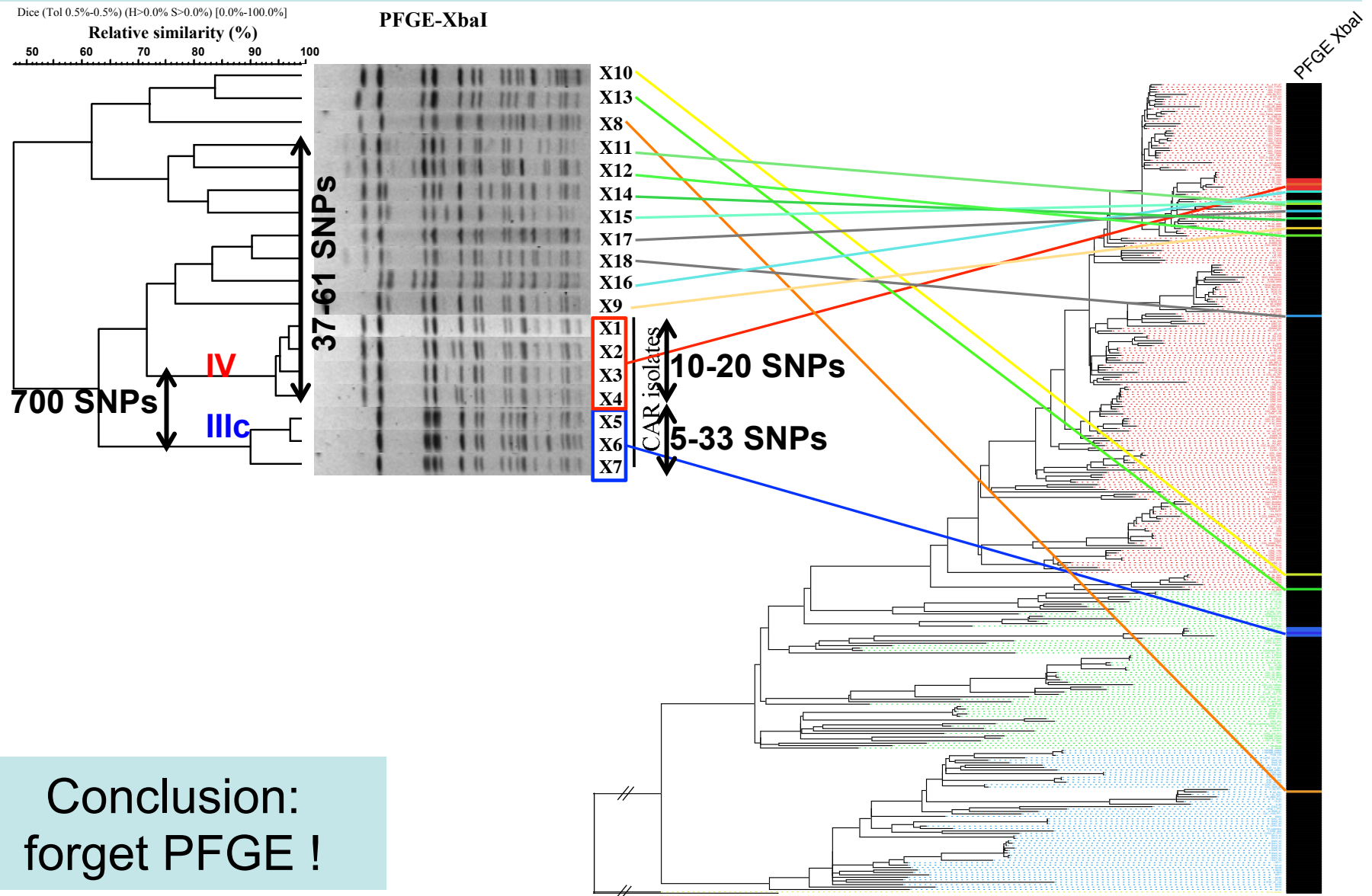
# Variants of the 66 kb SRL-PAI



- does not encode new resistance compared with plasmids
- *fec* operon
- better fitness ?



# Comparison of PFGE and whole genome sequencing



Conclusion:  
forget PFGE !

ML tree based on 14,677 SNPs

## Conclusions

- Sd1 is not a recent pathogen (MRCA 1757, 95% credible interval, 1645 – 1822) possibly linked to historical outbreaks of the 18th and 19th centuries
- Interest of historical isolates (extinct lineages, global understanding,...)
- Recent lineage IV from South Asia with iterative transmissions to Africa (such as *V. cholerae* and *Salmonella enterica* serotype Typhi H58)
- WGS should be used for investigating Sd1 outbreaks but not PFGE
- Accumulation of chromosomally-encoded multidrug resistance elements (waiting for NDM-1 and other carbapenemases)
- Efficient vaccine needed (WRSd1 IIIb, Sc-599 IIIc)

# Acknowledgements

S Brisse (PF8/GEB)

G Guigon (PF1)

E Souche (PF5)

C Bouchier (PF1)

C Bizet, D Clermont (CIP)

P Sansonetti (MMP)



E Njamkepo,  
N Fawal,  
A Tran-Dien,  
I Carle,  
M Lejay-Colin,  
C Ruckly,  
M Accou,  
M Gouali,  
S Le Hello,  
F Grimont,  
PAD Grimont

## RIIP

A Gassama (IPD)

MC Fonkoua (CPC)

R Bercion (IPB)

L Kaftyreva, S Egorova,

M Marekova (IPSP)



A Mather, A Page, M Aslett, S  
Harris, J Parkhill, N Thomson

epicentre AL Page,  
C Langendorf

CHU Saint-Pierre O Vandenberg



AL Wester



E Denamur

Public Health England C Jenkins, J Russell



H Izumya



C Jernberg

CDC N Strockbine, A Karas



M Venkatesan



R Murthy

icddr,b K Talukder



CS Chiou



MJ Blaser,  
G Perez-Perez

szu R Kolinska, M Marejkova



A Mérens



C Ye



A Ezernitchi, H Bercovier

K Kuleishov, A Podkolzin,  
G Schipulin



C Mason, L Bodhidatta



More information in



## **Global phylogeography and evolutionary history of *Shigella dysenteriae* type 1**

François-Xavier Weill *et al.*\*

And thank you for your attention !