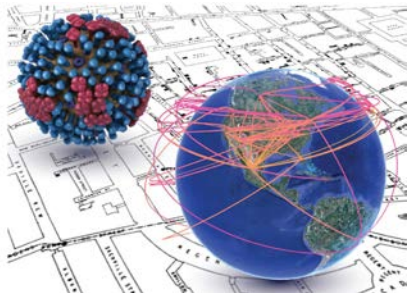


# Genomic approaches to reconstruct the landscape of microbial contamination in aquatic systems

## GLOBAL WATER FOOD SAFETY SUMMIT

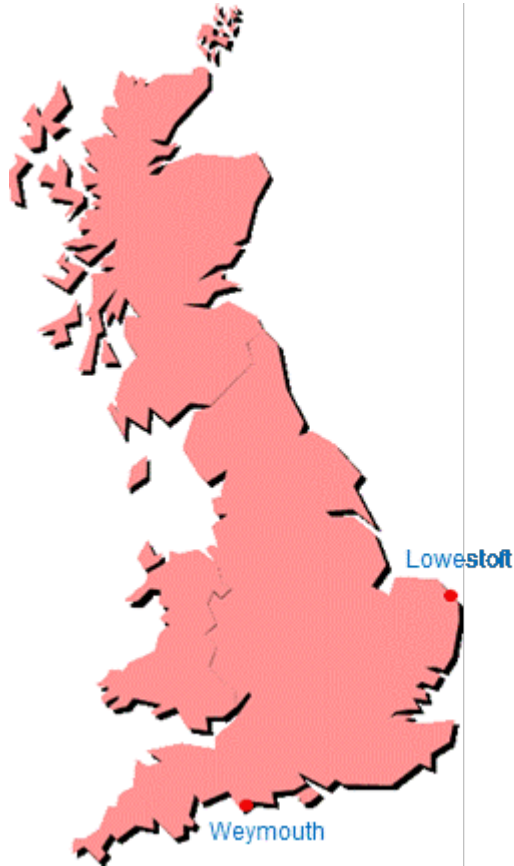


Prof. Jaime Martinez-Urtaza  
[jaime.martinez-urtaza@cefas.co.uk](mailto:jaime.martinez-urtaza@cefas.co.uk)  
[j.martinez-urtaza@exeter.ac.uk](mailto:j.martinez-urtaza@exeter.ac.uk)



# Cefas is an executive agency within the UK Government.

We currently employ over 500 staff at two main specialist laboratories and operate our own ocean going research vessel.



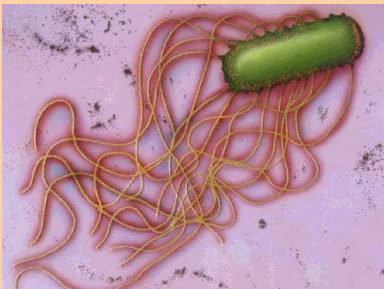
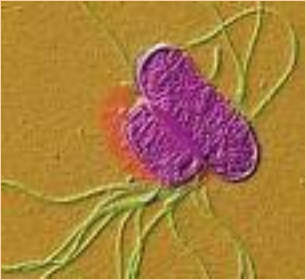
**Lowestoft:** international fisheries science and management, coastal and marine ecosystem processes, environmental radioactivity assessment, contaminant analysis, regulatory advice for coastal activities and information technology services

**Weymouth:** research and advice for fish health, disease diagnosis, fish and shellfish hygiene/food safety, ecotoxicology, evaluation of products used in aquaculture

# Populations dynamics and diseases

## Two Models:

*Salmonella*

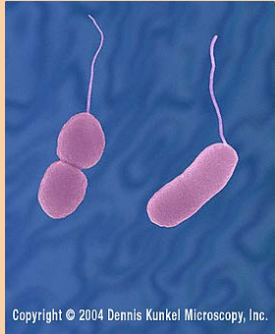


Foodborne pathogen

*Vibrio*



*V. cholerae*



*V. parahaemolyticus*

Waterborne pathogen

Classic Typing  
Techniques



Whole Genome  
Sequencing

**MLVA**

**PFGE**

**MLST**



**Genomic  
Epidemiology**



## Core genome

- Mutations
- Recombination



## Phylogeny



## Accessory Genome

- Horizontal Gene

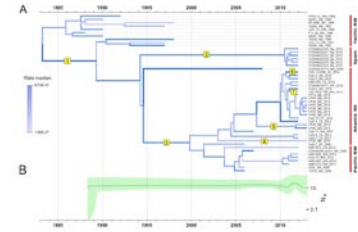
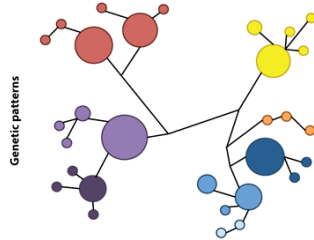
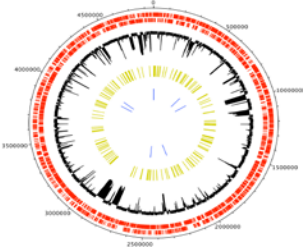
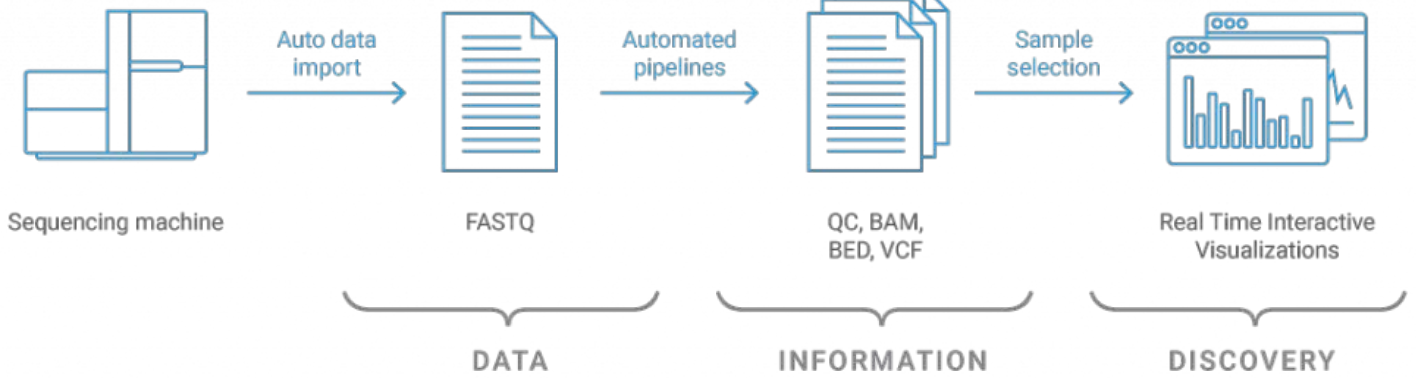
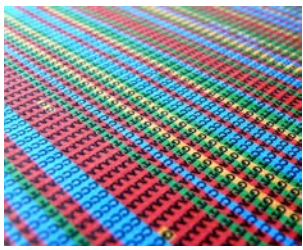


## Local Adaptation

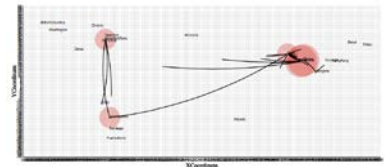
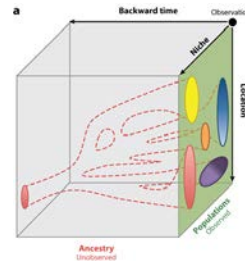
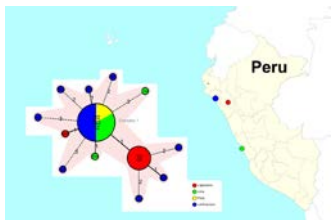
AB Resistance

Virulence

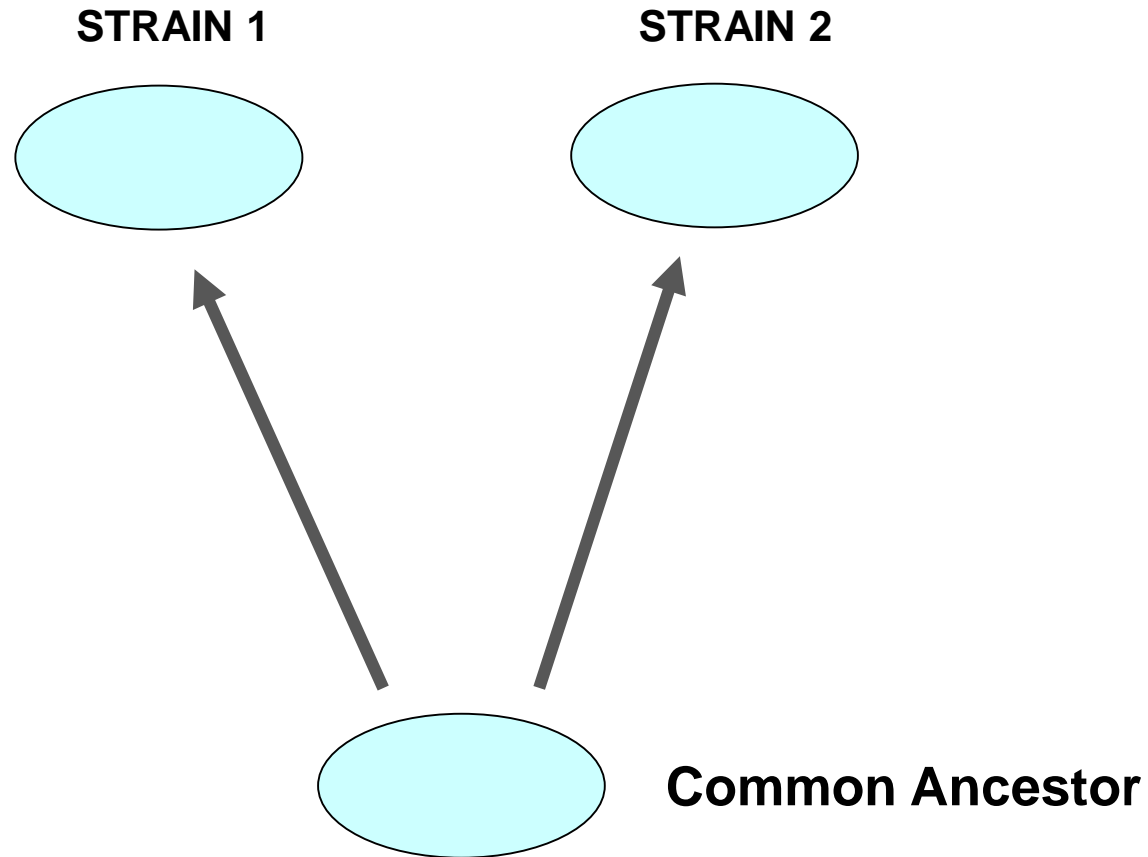








Genome → Population → Demography →



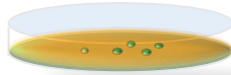
Biogeography → Evolution → Colonization



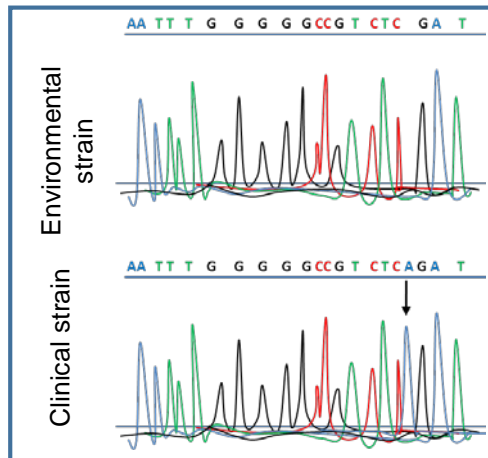
- When?  Date of emergence of clones and variants
- Where?  Place of emergence and routes of dispersal
- Introduction?  Route of entry for the pathogen
- Adaptation?  Likelihood of becoming endemic

# Biological 'corridors' of disease?

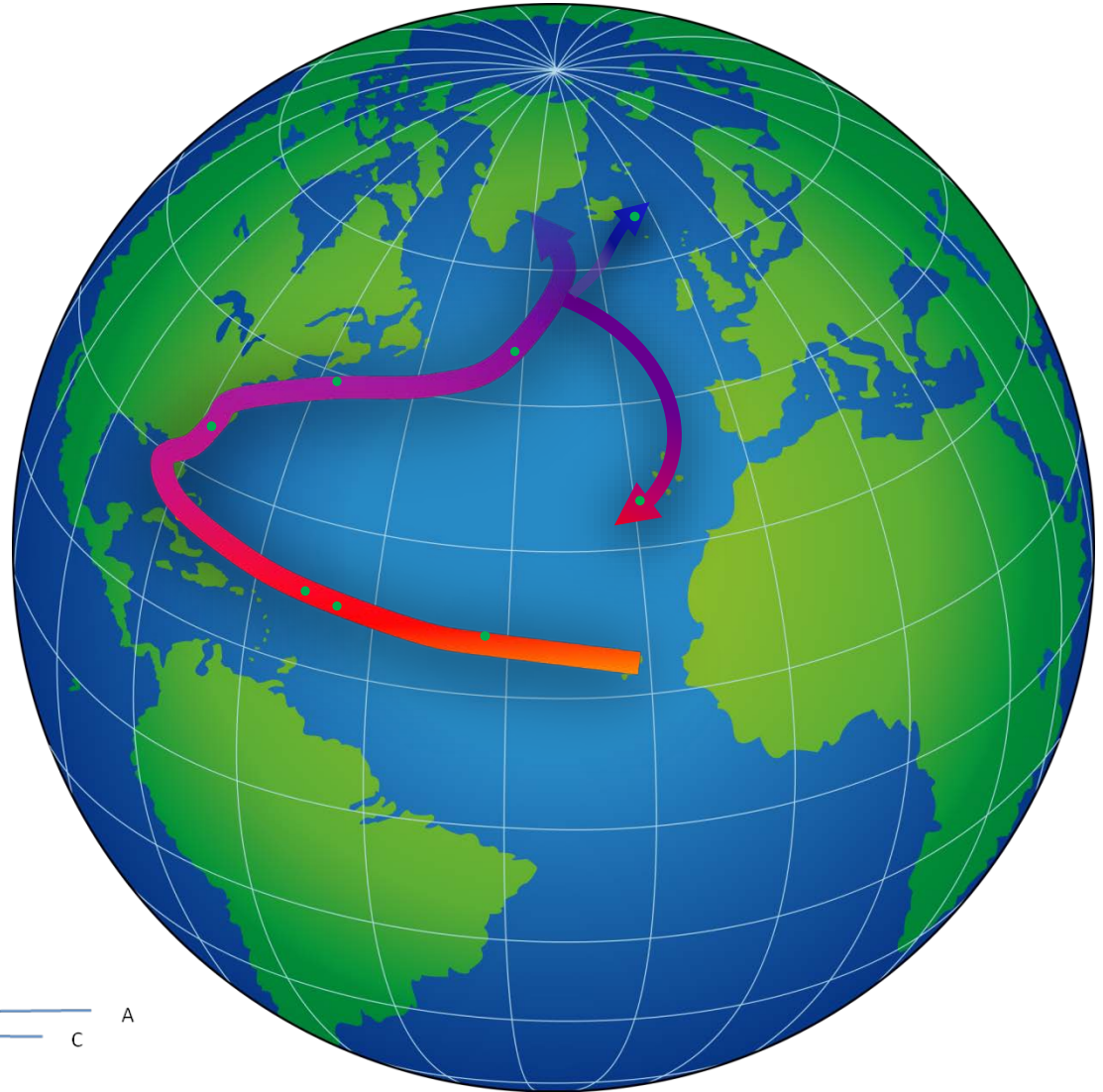
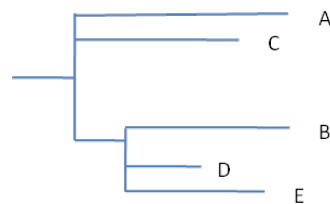
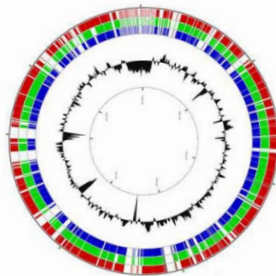
1) Isolation of (•) strains



2) Genome sequencing and SNP identification

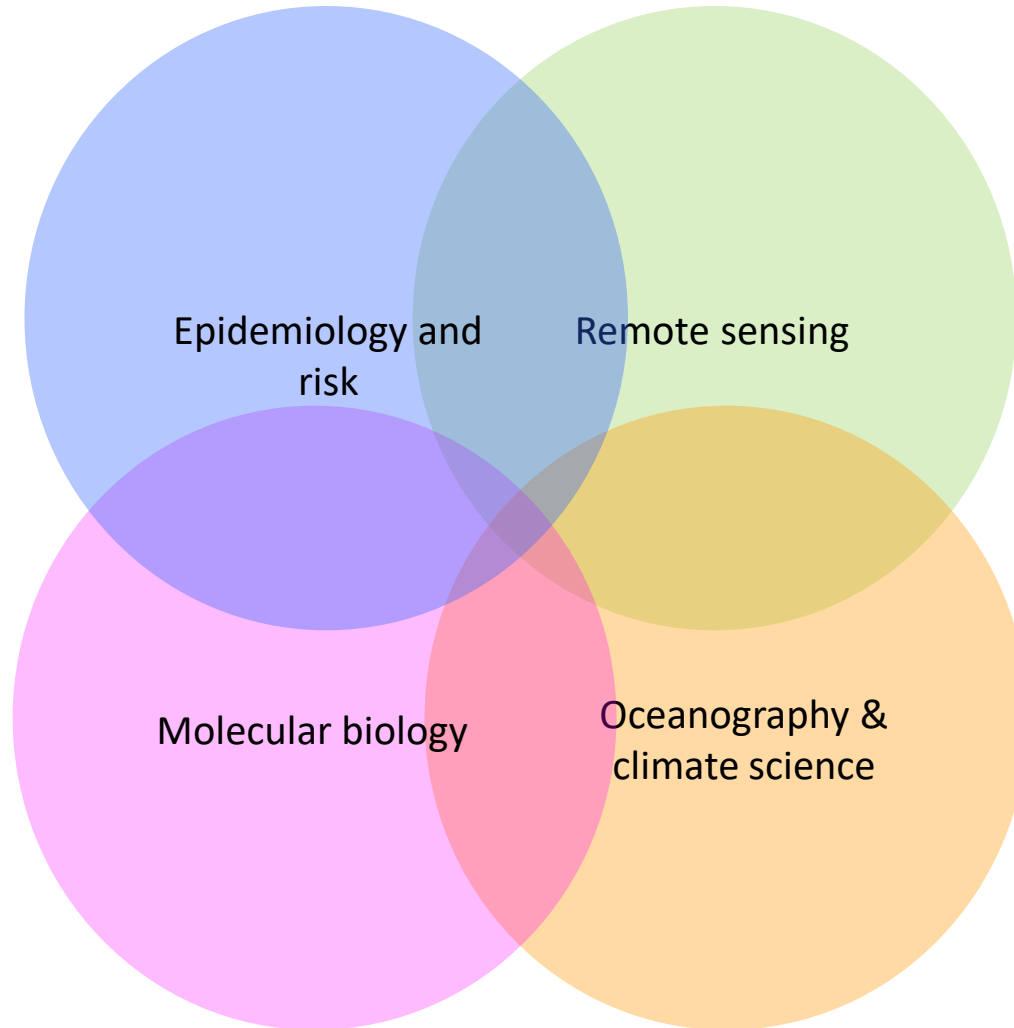


3) Comparative sequence analysis and phylogenetic determination





# Analysing disease emergence using a combination of approaches



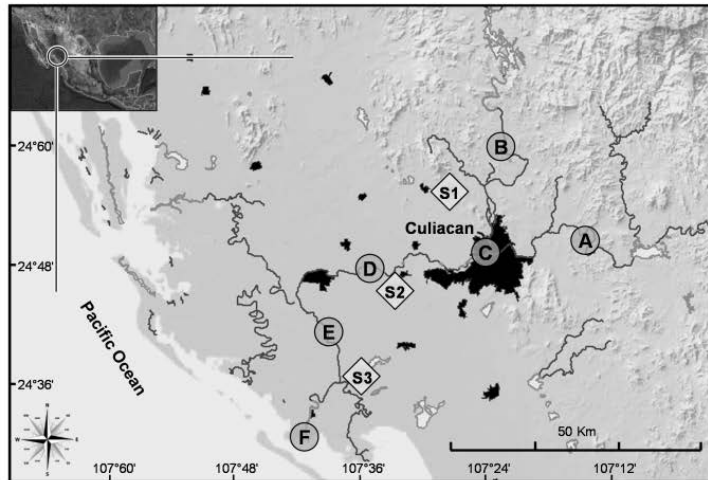


# Application of genome-wide analysis for tracking the dispersion of *Salmonella enterica* strains in rivers of Sinaloa, Mexico

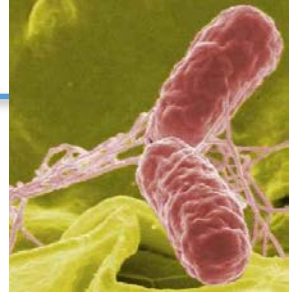
J. R. Ibarra-Rodríguez, C. Chaidez, J. R. Aguirre-Sánchez, I. F. Vega-López, M. Hoffmann, N. Gonzalez-Escalona, J. Martinez-Urtaza

*Salmonella*

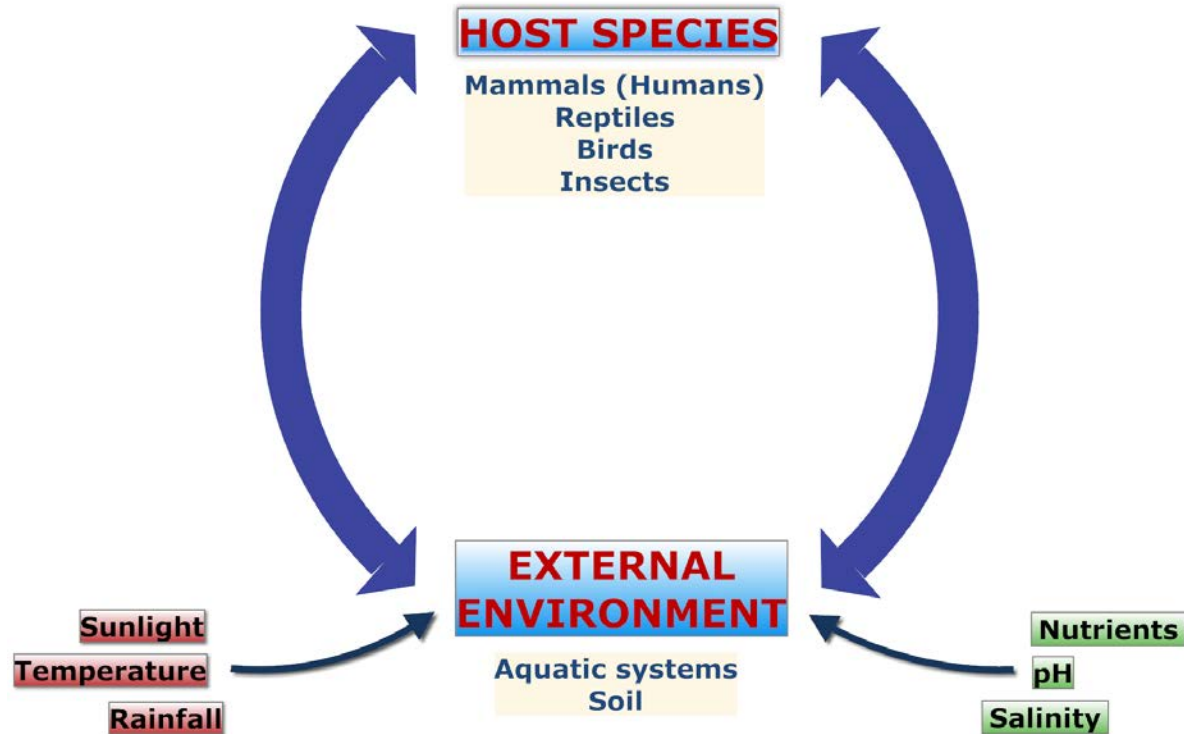
## Phylodynamics of contamination and transmission of *Salmonella* in aquatic systems



# Salmonella



- Pathogen of human and animals
- One of the major causes of foodborne infection at global scale



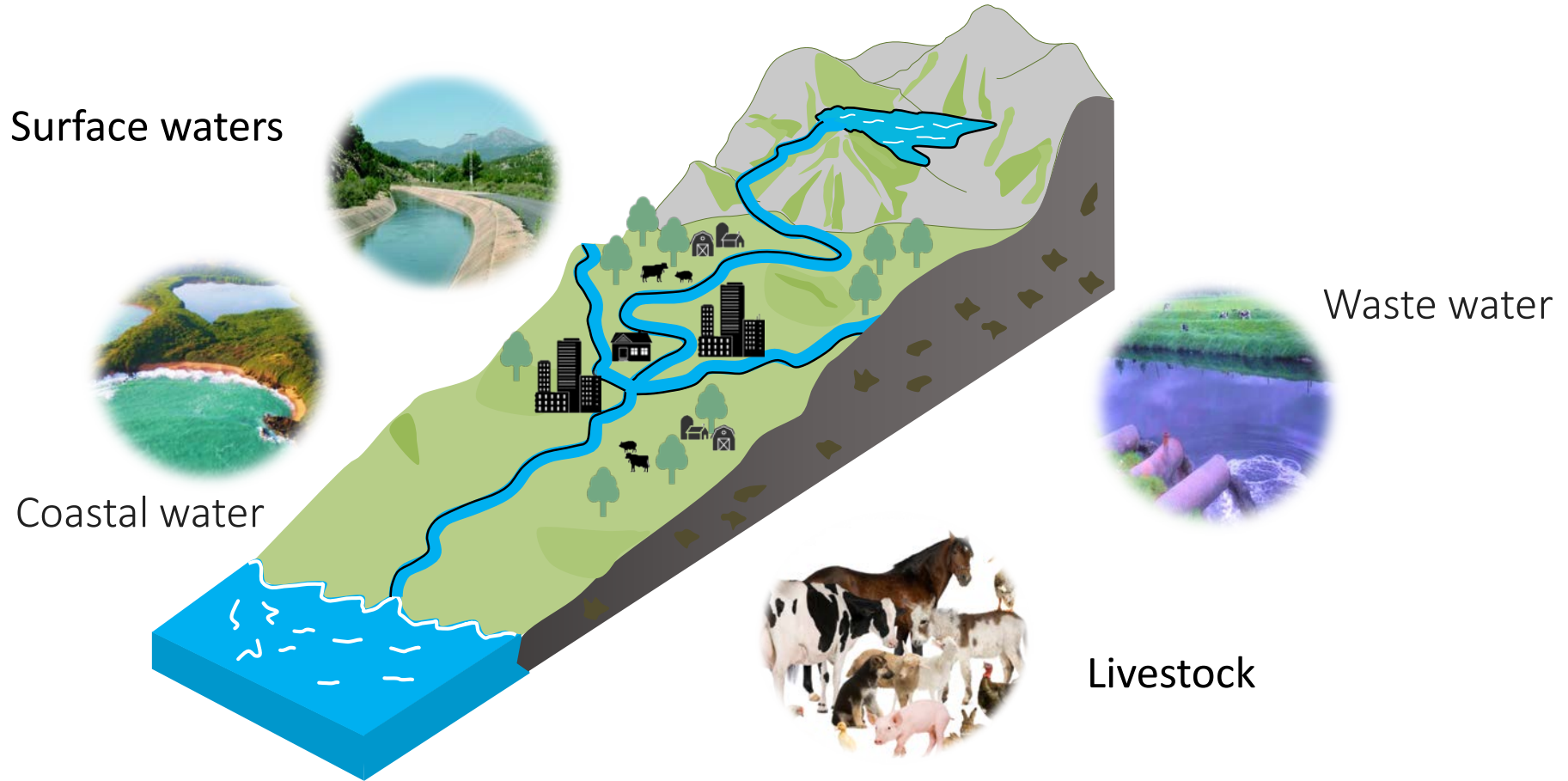
# Sinaloa, Mexico



- Mexico is characterized by a **high morbidity** associated with salmonellosis.
- 119,374 cases in 2010, representing **106 cases per 100,000 inhabitants**
- Sinaloa, nearly twice the number of cases, **192 cases per 100,000** inhabitants
- United States, **16.7 cases per every 100,000** population

# Sinaloa

## Phylodynamics of Salmonella contamination in the river Culiacán







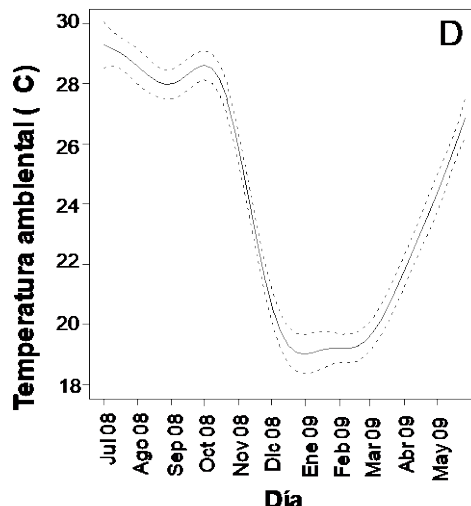
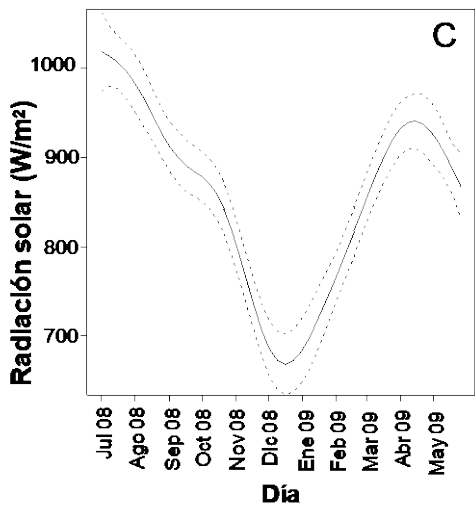
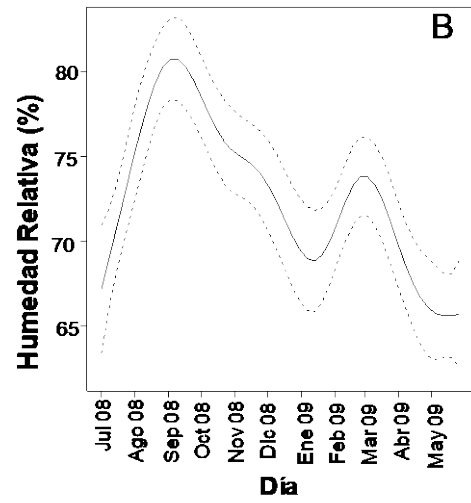
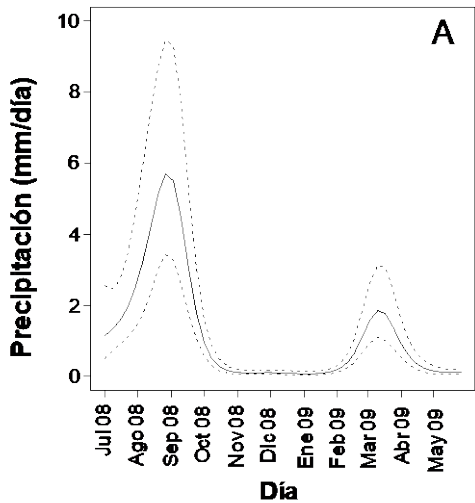
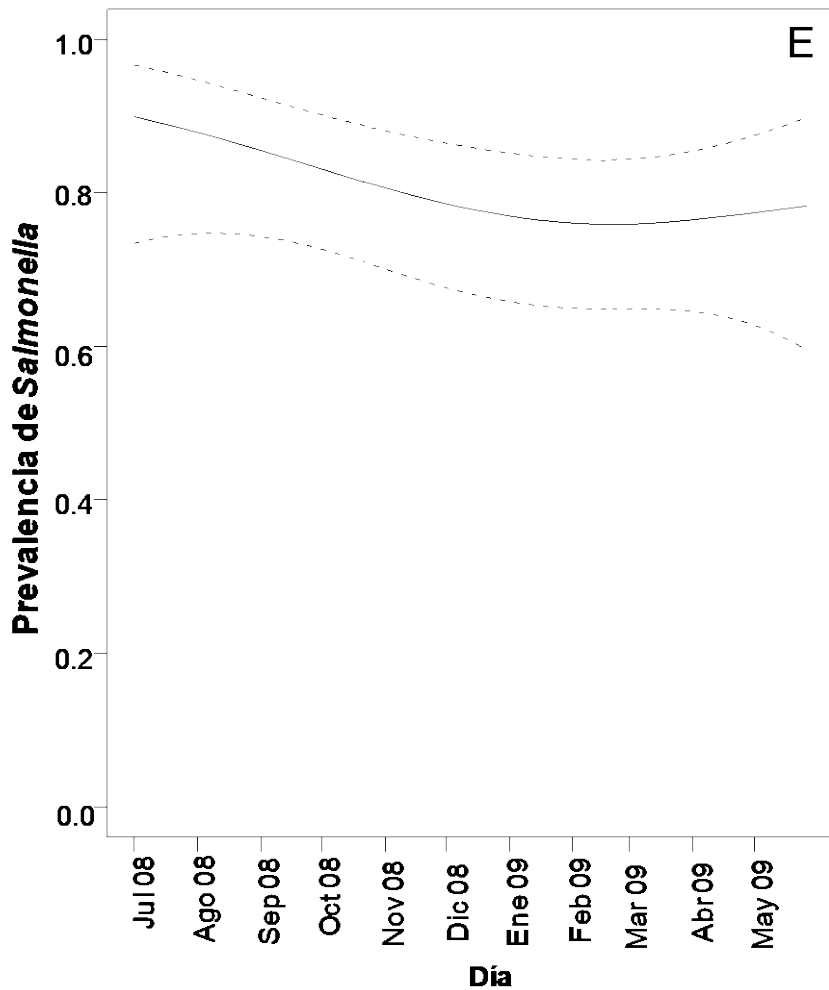
- Despite the growing importance of non-host ecology in resolving the epidemic dynamics of Salmonella, investigations of this pathogen in natural settings has been typically constrained by the ***low occurrence of this organism in environmental samples***, typically below ***5 %***.
- The rare presence of Salmonella in the environment has made extremely difficult the identification of patterns of contamination in a specific region or ecosystem, ***limiting any possibility for a reliable traceback*** from original sources which ultimately allows for the delineation of the routes of dispersal.
- The low number of strains identified over the course of environmental surveys typically belong to a ***broad range of serovars with highly diverse genetic backgrounds***.
- This situation limits the comparison of strains at a serovar level and any effort to ***compare strains of different serovars would result inconclusive***.

# Sinaloa, Mexico

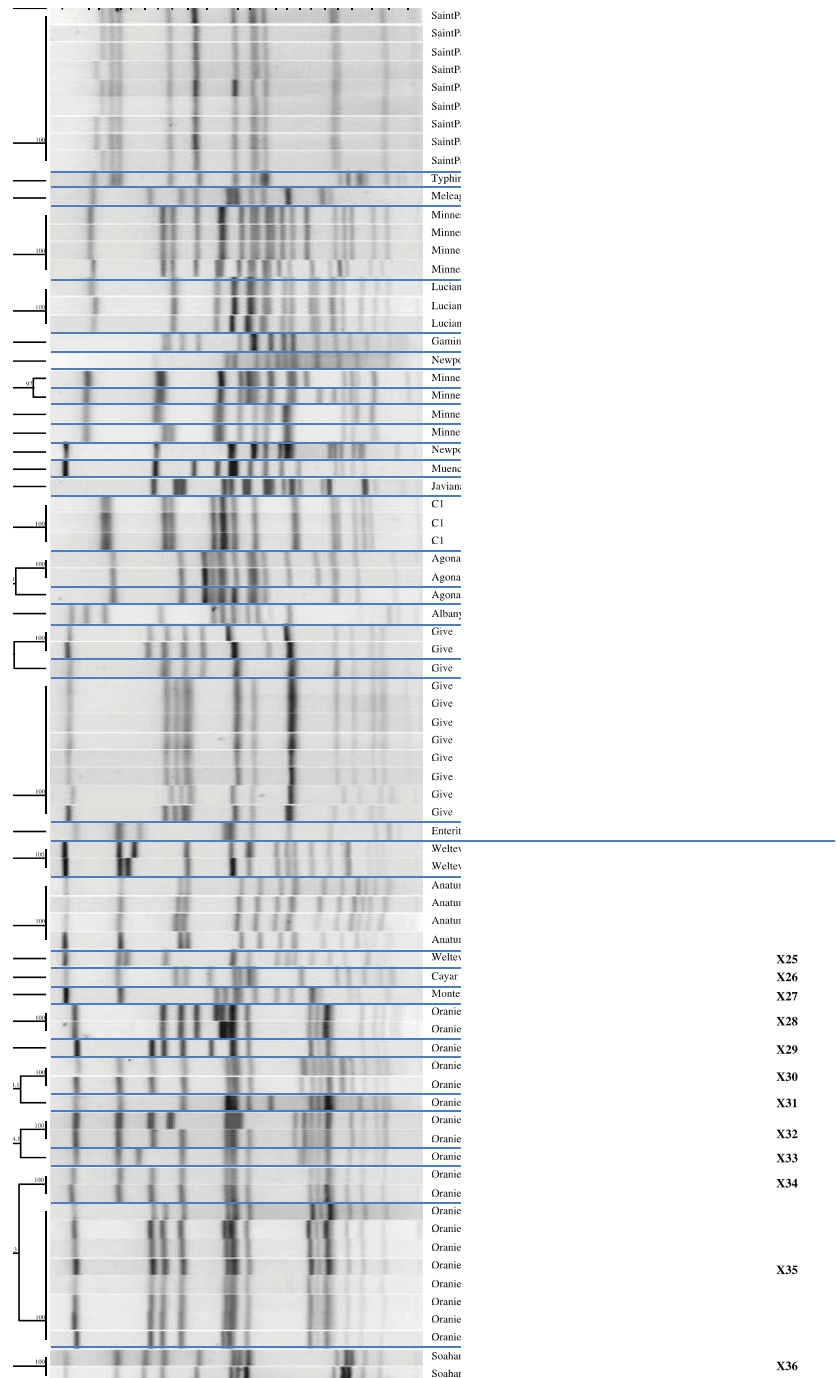


Location	Prevalence (%)	IC (95%)
Jotagua	95.65	(82.22, 99.75)
Aguacaliente	69.56	(49.39, 85.60)
Puente negro	86.95	(69.57, 96.59)
Cofradía de Navolato	78.26	(58.96, 91.60)
Iraguato	91.30	(75.50, 98.50)
El castillo	60.87	(40.56, 78.88)

# Tropical Climate



# Pulsed-field gel electrophoresis PFGE



Microb Ecol (2011) 61:811–820  
DOI 10.1007/s00248-010-9792-5

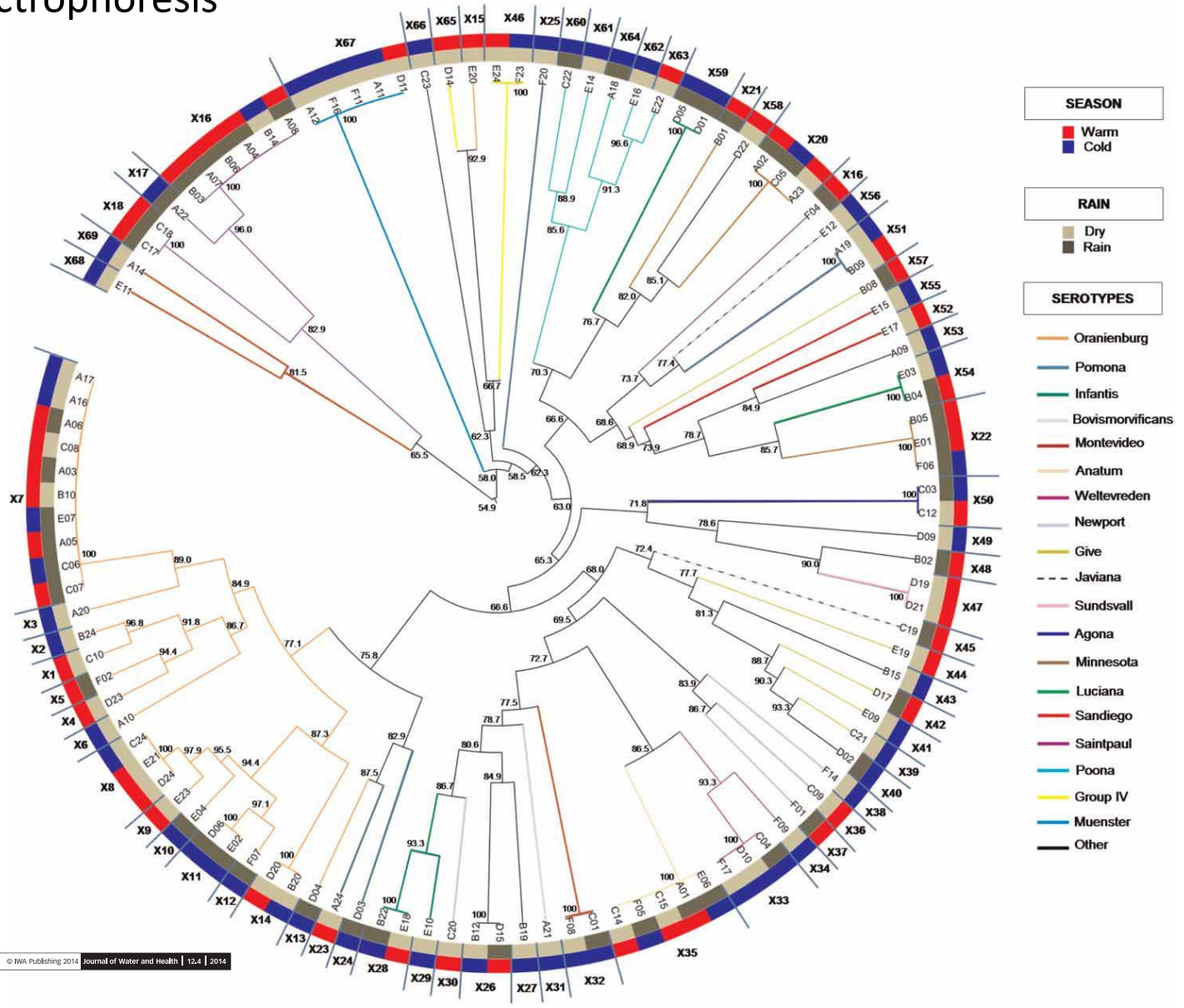
ENVIRONMENTAL MICROBIOLOGY

## Geographical and Temporal Dissemination of Salmonellae Isolated from Domestic Animal Hosts in the Culiacan Valley, Mexico

Maribel Jiménez · Jaime Martínez-Urtaza · Cristóbal Chaidez

# Pulsed-field gel electrophoresis

# PFGE



**SEASON**

Warm  
Cold

**RAIN**

Dry  
Rain

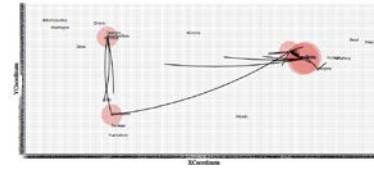
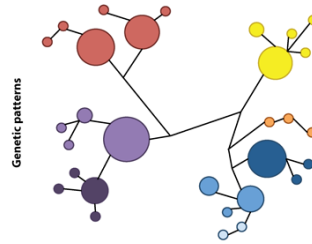
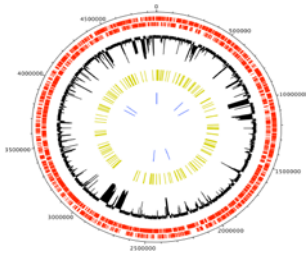
**SEROTYPES**

- Oranienburg
- Pomona
- Infantis
- Bovismorificans
- Montevideo
- Anatum
- Welltevreden
- Newport
- Give
- Javiana
- Sundsvall
- Agona
- Minnesota
- Luciana
- Sandiego
- Saintpaul
- Poona
- Muenster
- Other

## Prevalence and genetic diversity of *Salmonella* spp. in a river in a tropical environment in Mexico

Maribel Jiménez, Jaime Martínez-Urtaza, María Xosé Rodríguez-Alvarez, Josefina León-Félix and Cristóbal Cháidez





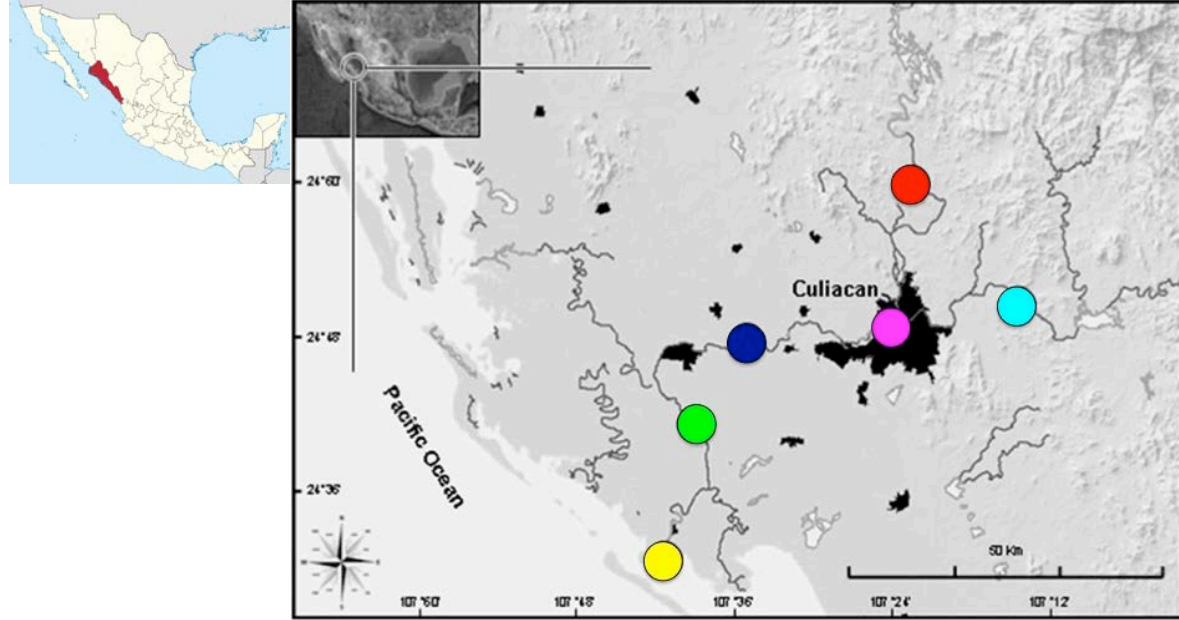
**AIM:** to test the efficiency of WGS to identify the point source of contamination and draw the routes of dispersal of Salmonella in a specific geographic area (Sinaloa, Mexico).

??

➔ Are these populations resident in the river?

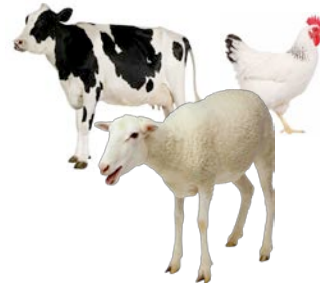
➔ Could we use genomics to track the dispersion?

# The Study



- A total of **150** *Salmonella enterica* strains (Jiménez et al., 2011, 2014) were sequenced. Only **61** local *Salmonella enterica* genomes belonging to the two prevailing serovars -Oranienburg and Saintpaul- were selected for the study
- **Six sampling sites** were selected to cover the study area (named A, B, C, D, E and F). Sites A and B were located on mountain sides; Site C was located in Culiacan City; Sites D and E were located in the valley next to the city limit; and Site F was located on the Pacific coast.
- Strains were isolated from **river water** and **animal feces** (cow, goat and chicken)

# Strains



Serotype	Total
Oranienburg	19
Give	11
Saintpaul	9
Minnesota	8
Anatum	4
Luciana	3
Thompson	3
Weltevreden	3
Agona	3
Newport	2
Soahanina	2
Albany	1
Montevideo	1
Muenchen	1
Gaminara	1
Cayar	1
Javiana	1
Meleagridis	1
Enteritidis	1
Typhimurium	1
<b>Total</b>	<b>76</b>

Serotype	Total
Oranienburg	23
Saintpaul	10
Minnesota	7
Give	5
Anatum	5
Muenster	5
Pomona	5
Infantis	5
Poona	5
Weltevreden	4
Montevideo	4
Newport	3
Bovismorbificans	2
Sandiego	2
Javiana	2
Sundsvall	2
Luciana	2
Agona	2
Seftenberg	1
Braenderup	1
Kiambu	1
Soahanina	1
Texas	1
<b>Total</b>	<b>98</b>

Country	Strains
Australia	2
Brazil	2
China	3
Colombia	10
Costa Rica	2
Denmark	1
Egypt	1
Etiopia	1
Filipinas	2
France	4
Ghana	1
Honduras	1
India	26
Indonesia	2
Jamaica	1
Lebanon	1
Malaysia	2
Malta	1
Mexico	23
Saudi Arabia	1
Tailand	1
Turkey	8
Uk	80
USA	117
Vietnam	2
<b>Total</b>	<b>295</b>

# Sequencing strategy

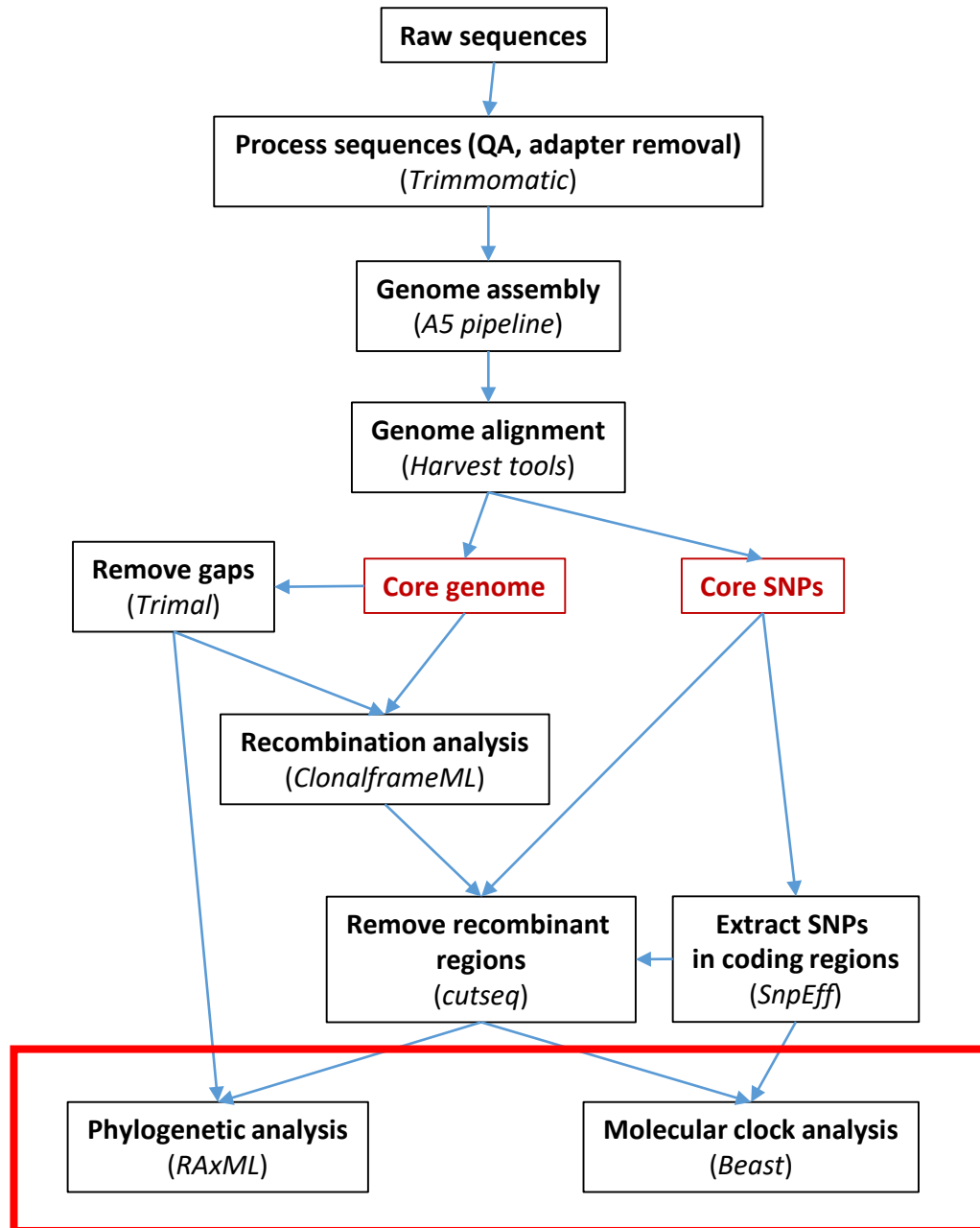
- Strains were sequenced by MiSeq (Illumina)
- Minimum coverage of 40–120X.
- Libraries were prepared with the Nextera XT DNA sample preparation kit (Illumina).



## PacBio (FDA)

- S\_Muenster\_CFSAN001301
- S\_Give\_CFSAN024229
- S\_Poona\_ATCC\_BAA\_1673
- S\_Infantis\_CFSAN003307
- S\_Oranienburg\_CFSAN001285
- S\_Saintpaul\_CFSAN004173
- S\_Minnesota\_CFSAN017963
- S\_Pomona\_ATCC\_10729\_CFSAN000720



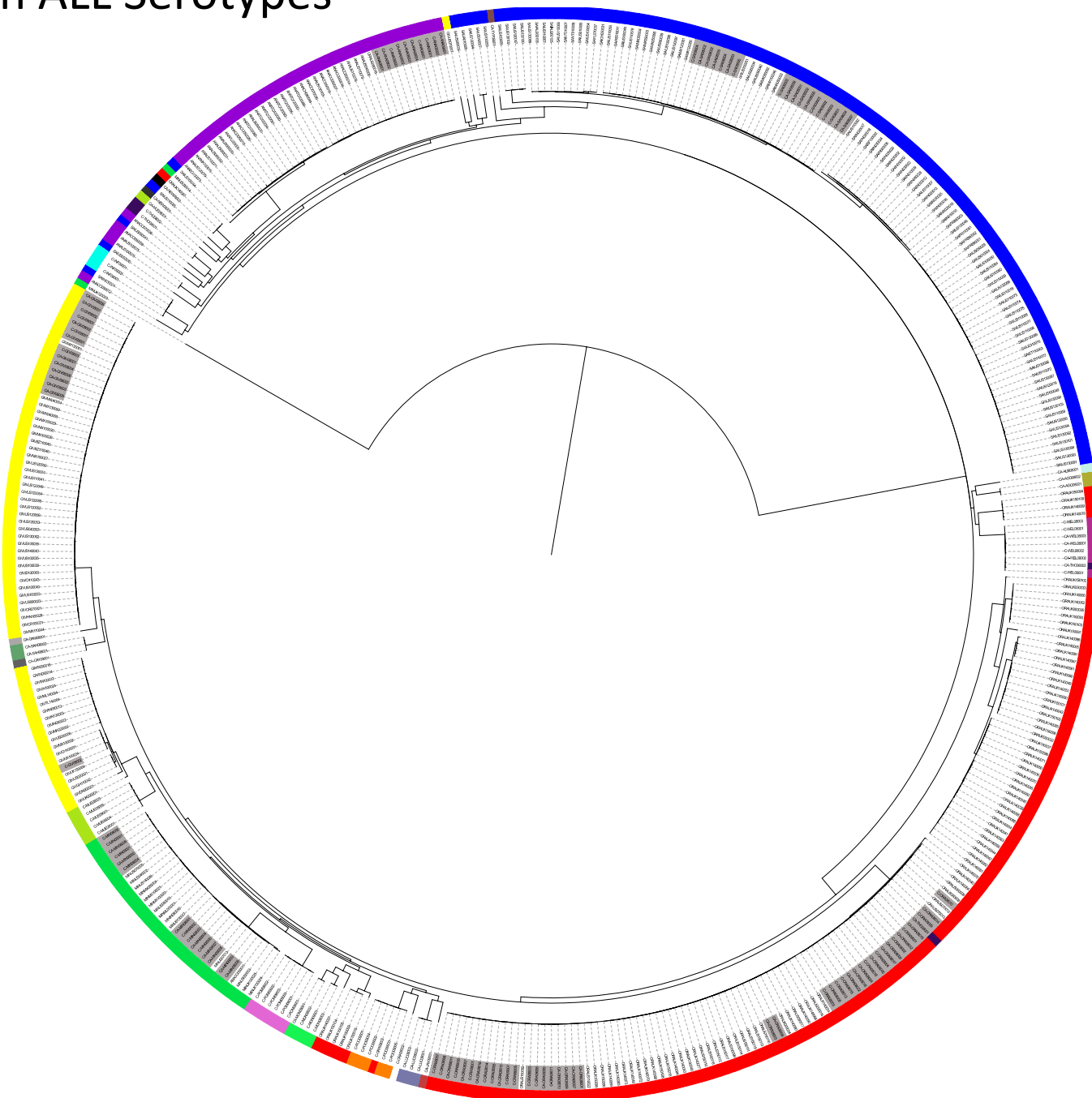




# Global phylogeny with ALL Serotypes

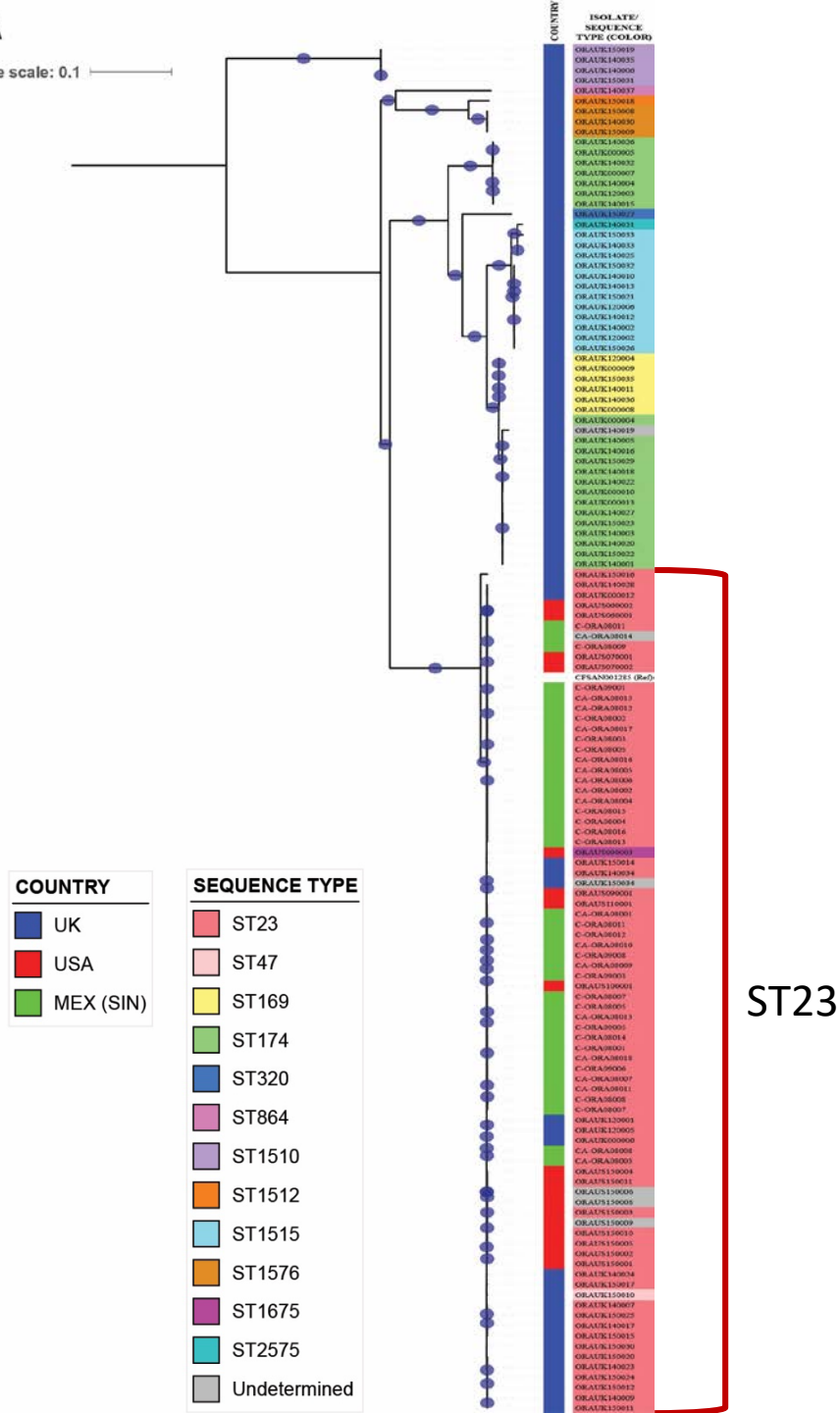
Tree scale: 0.1

- Global**
- Oranienburg
  - Saintpaul
  - Minnesota
  - Give
  - Anatum
  - Poona
  - Infantis
  - Thompson
  - Newport
  - Muenchester
  - Pomona
  - Agona
  - Weltevreden
  - Montevideo
  - Typhimurium
  - Luciana
  - Gaminara
  - Sahanina
  - Cayar
  - Albany



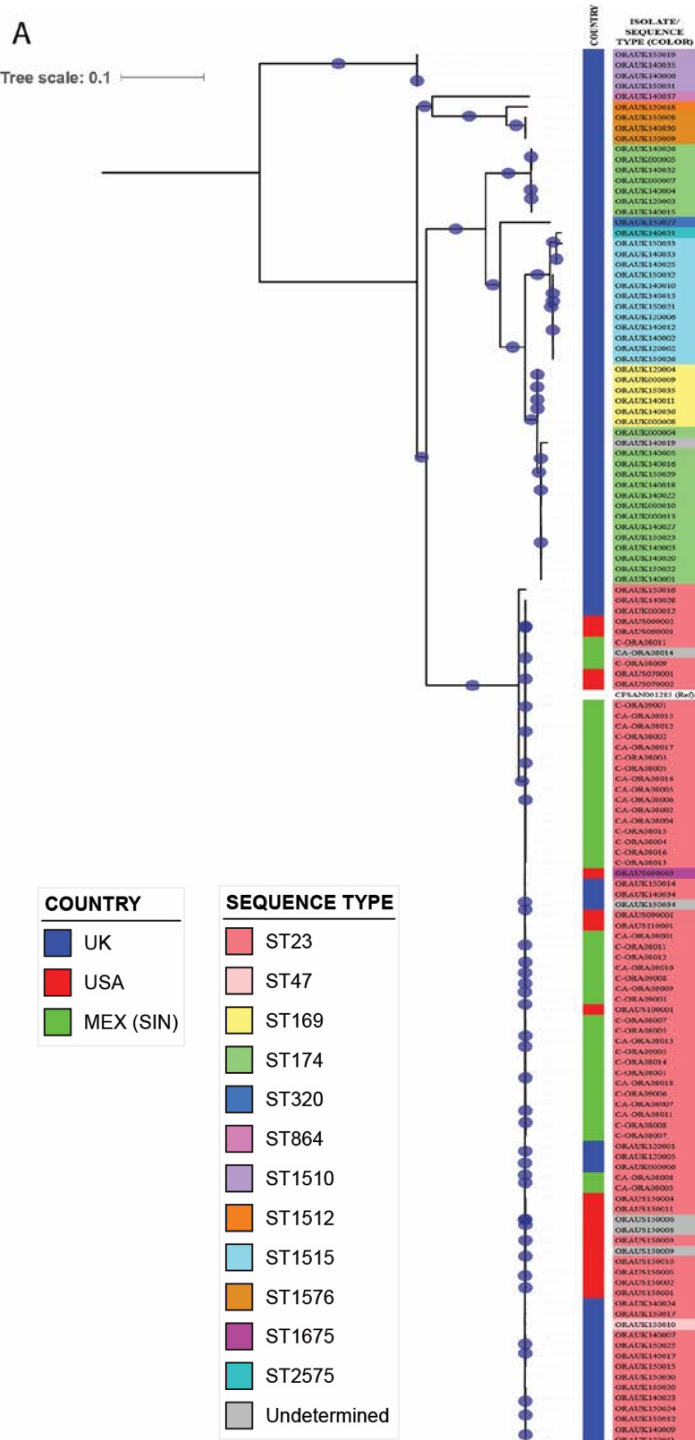
A

Tree scale: 0.1

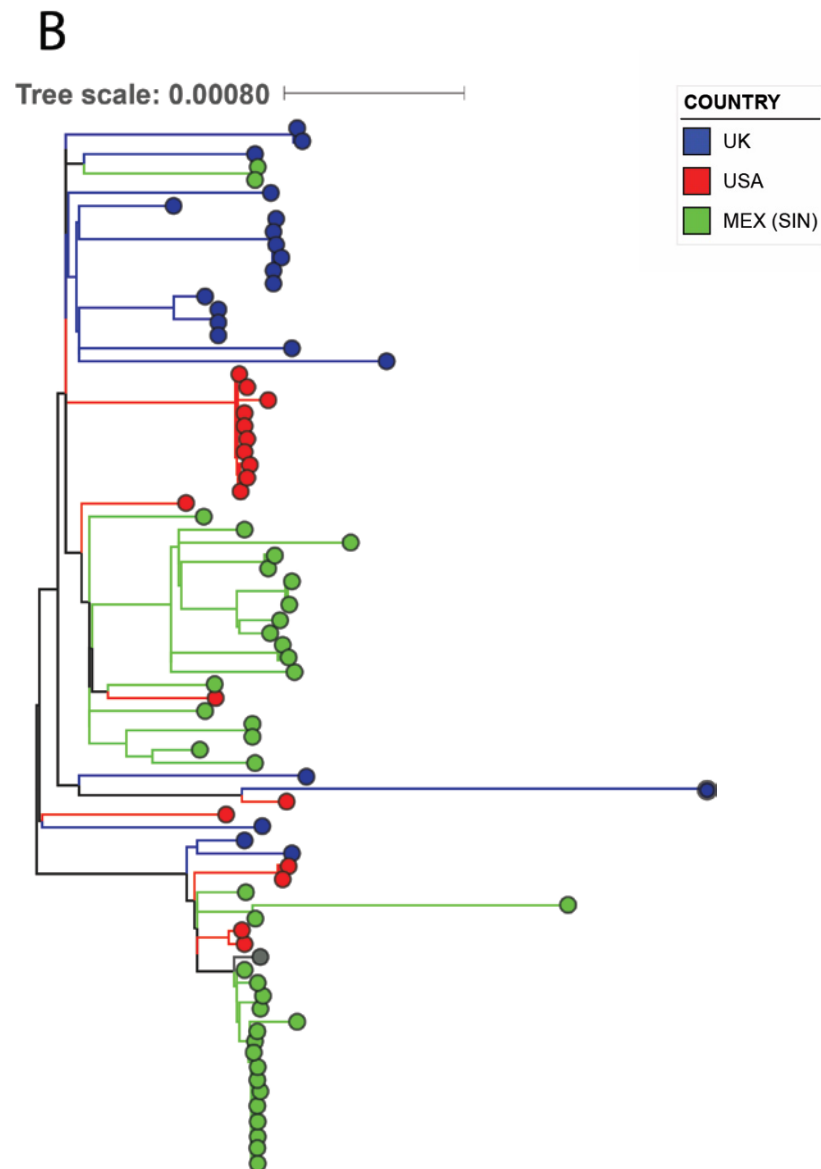


Global phylogeny of *Salmonella* Oranienburg. Maximum likelihood tree of the full dataset showing the phylogenetic relations between the core genome of 92 global and 40 local *S.* Oranienburg isolates.

ST23

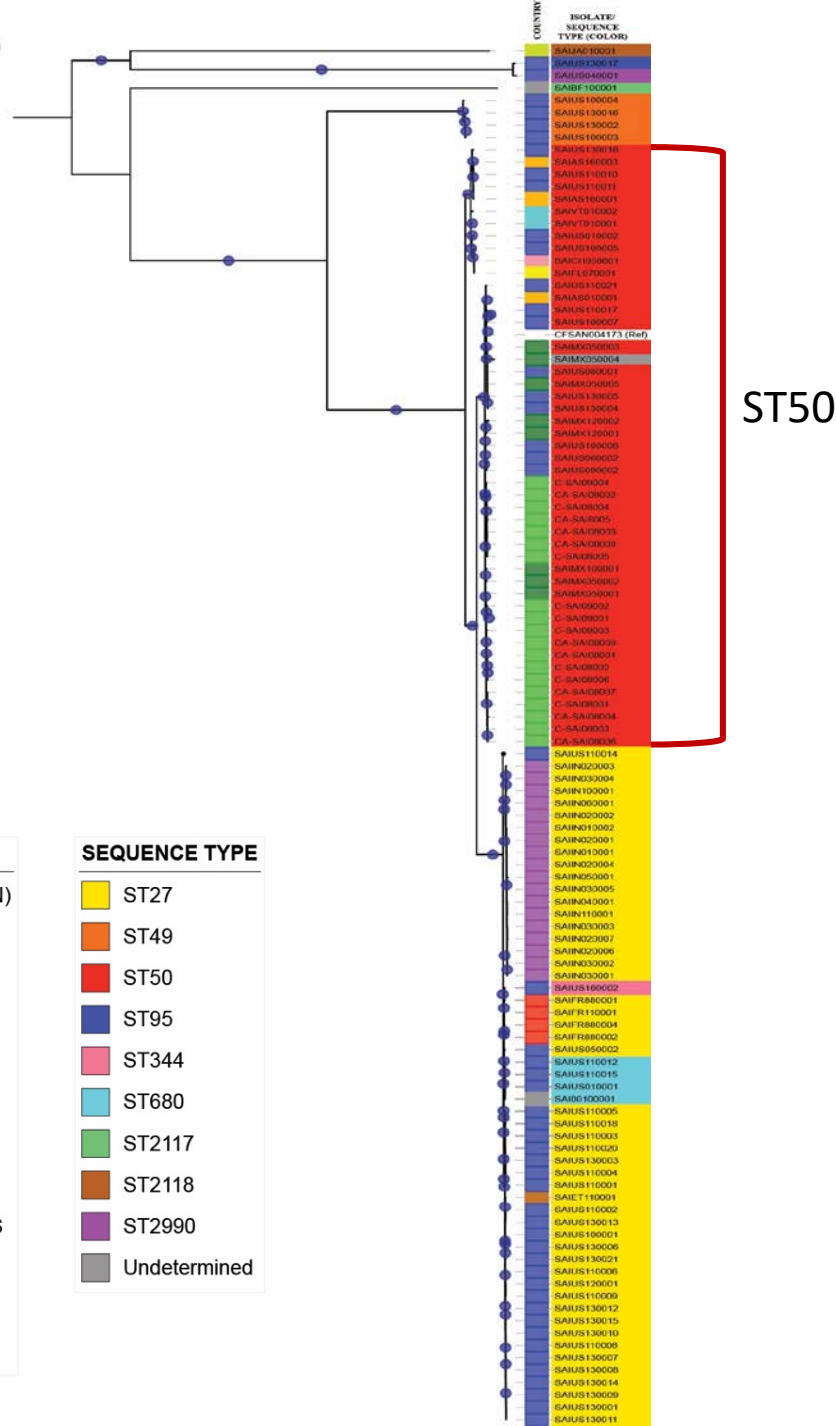


Phylogeny of *Salmonella* Oranienburg ST23 comprising all the isolates from Sinaloa.



A

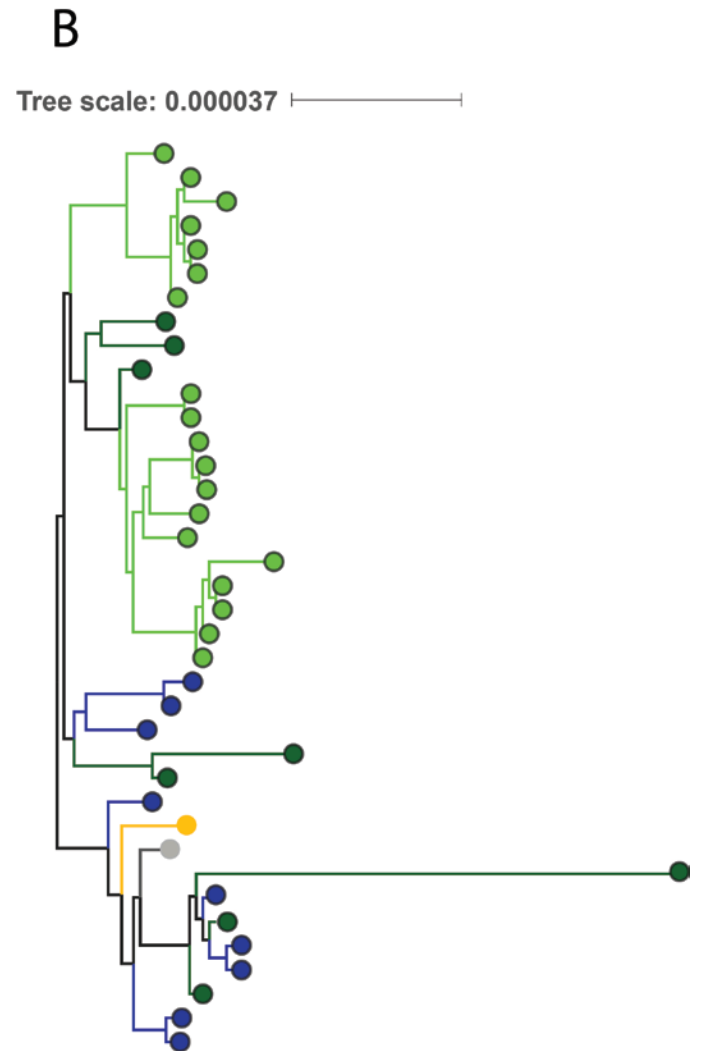
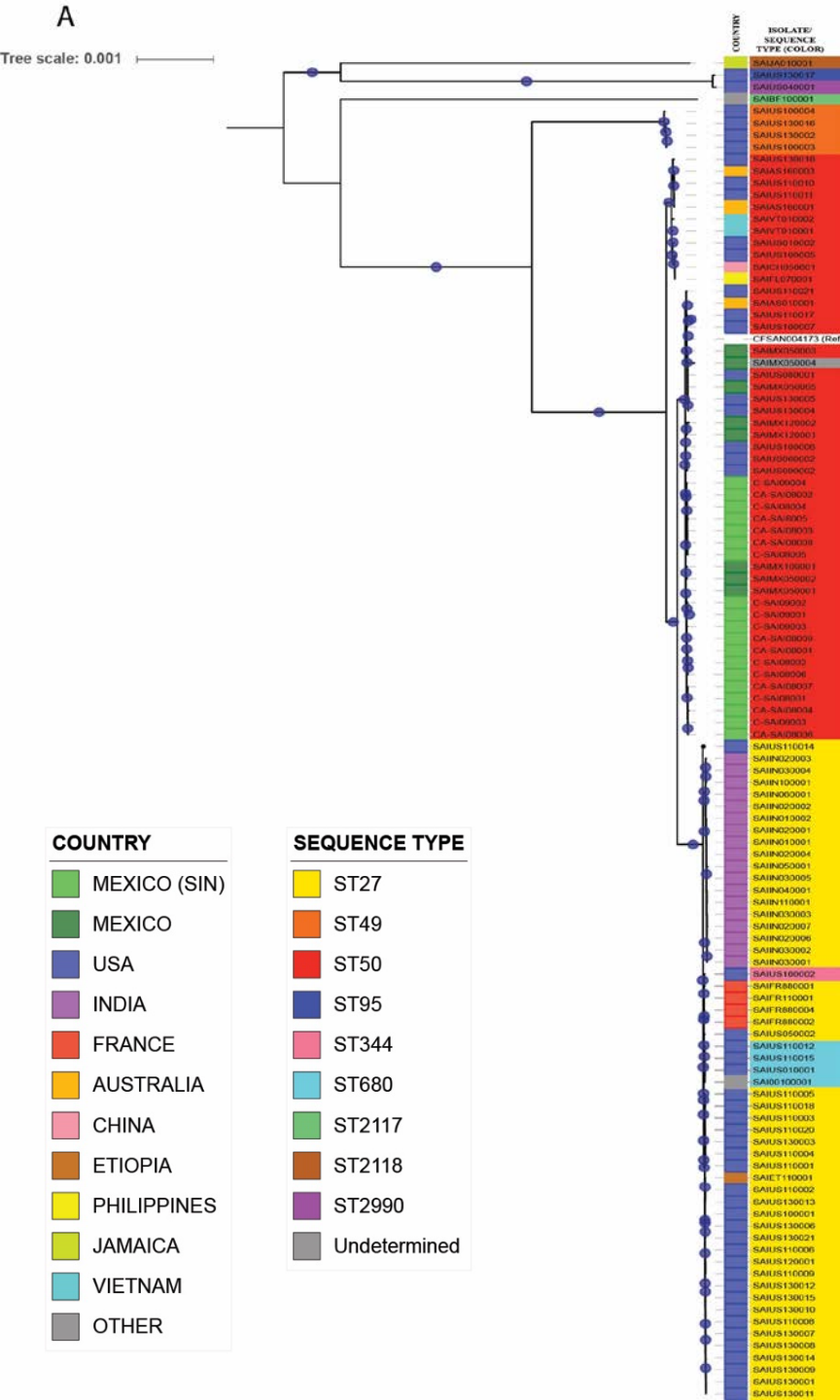
Tree scale: 0.001



Global phylogeny of *Salmonella* Saintpaul. Maximum likelihood tree showing the phylogenetic relations between the core genome of 93 global and 19 local isolates.

COUNTRY	SEQUENCE TYPE
MEXICO (SIN)	ST27
MEXICO	ST49
USA	ST50
INDIA	ST95
FRANCE	ST344
AUSTRALIA	ST680
CHINA	ST2117
ETIOPIA	ST2118
PHILIPPINES	ST2990
JAMAICA	Undetermined
VIETNAM	
OTHER	

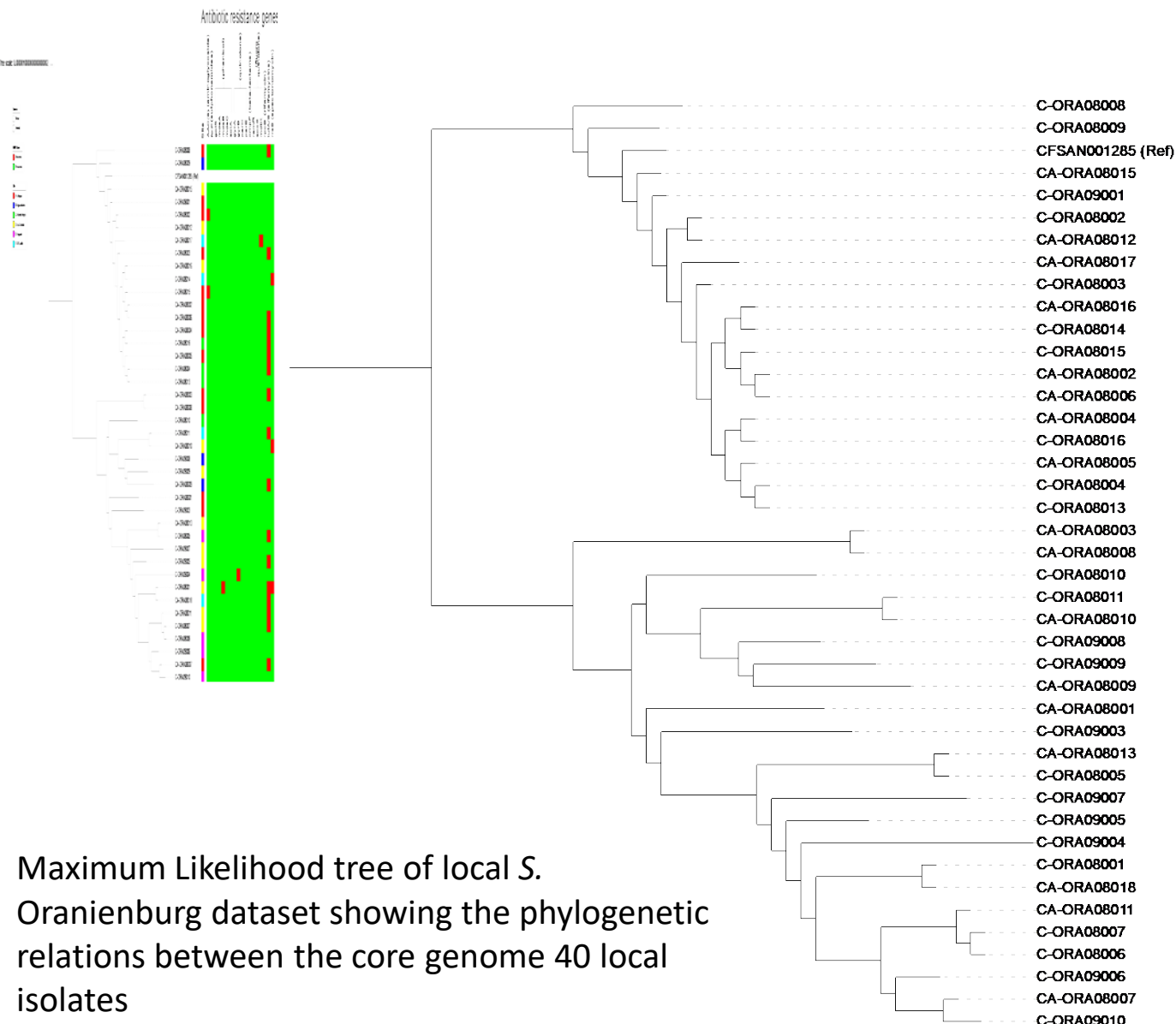
ST50



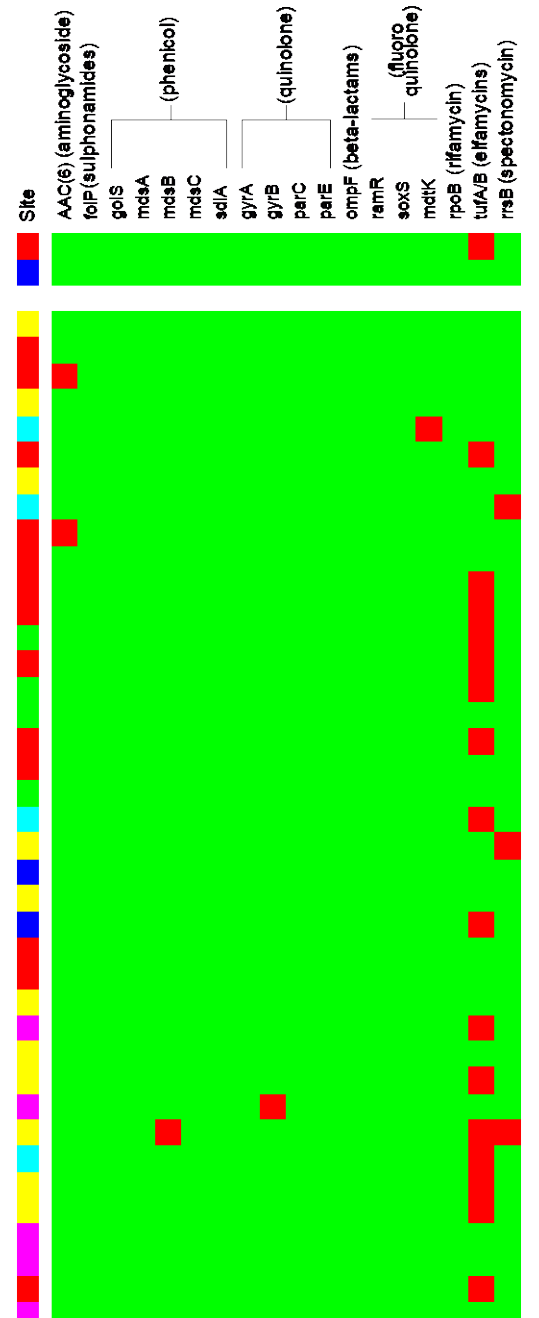
Phylogeny for the specific clade of ST50 where all the isolates from Sinaloa were included.

# Local phylogeny and resistome cluster analysis of *Salmonella* Oranienburg.

Tree scale: 0.000001000000000000000002

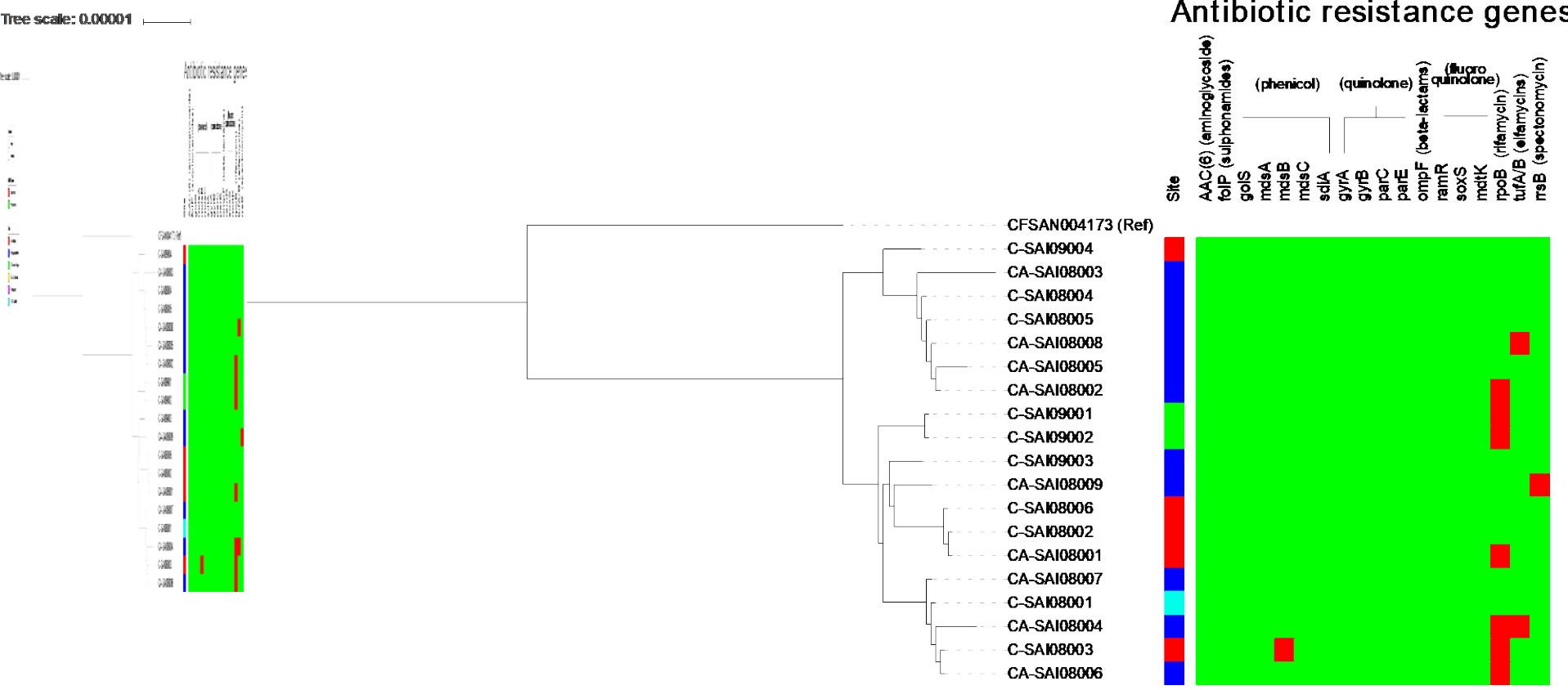


## Antibiotic resistance genes



Maximum Likelihood tree of local *S.* Oranienburg dataset showing the phylogenetic relations between the core genome 40 local isolates

# Local phylogeny and resistome cluster analysis of *Salmonella* Saintpaul.



Maximum Likelihood tree of local *S. Saintpaul* dataset showing the phylogenetic relations between the core genome 19 local isolates

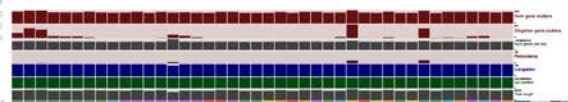
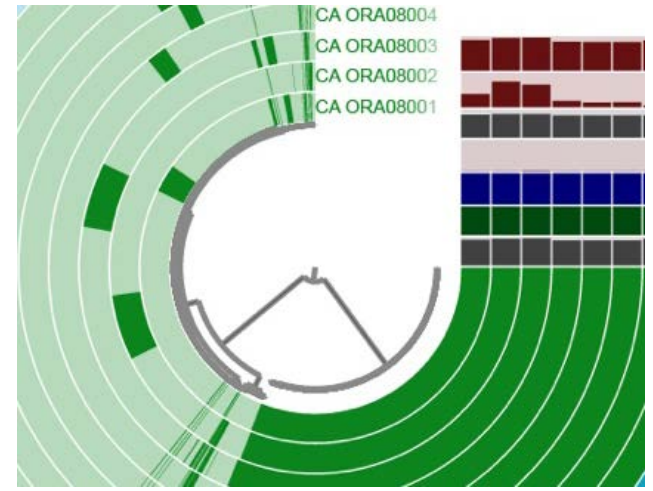
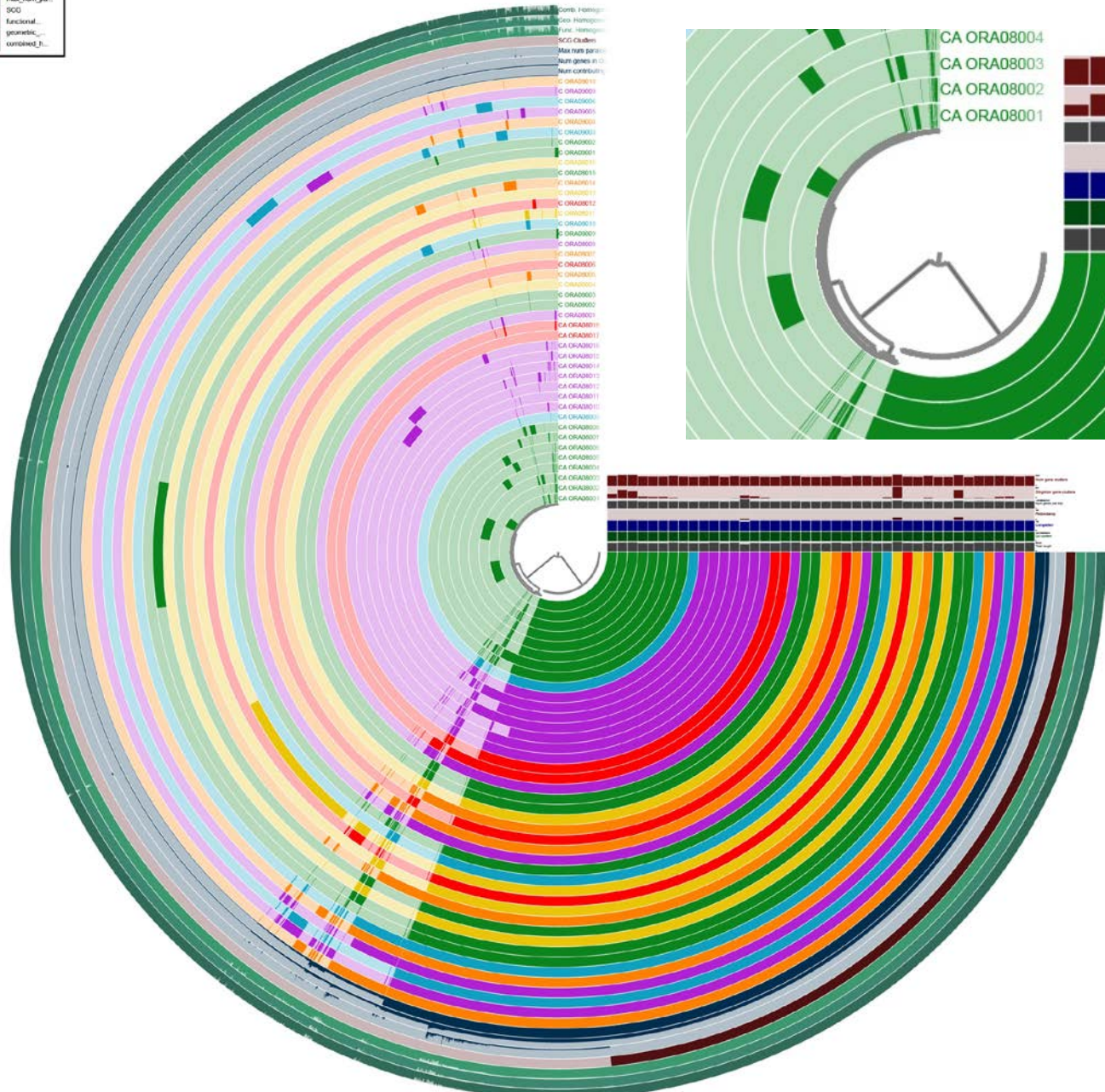


# ORANIENBURG PAN

Items order: Presence absence (D: Euclidean; L: Ward) | Current view: gene\_cluster\_presence\_absence | Samples order: custom

**Layers**

- Color keys
- Intagria
- Agua caliente
- Cofradia
- Cadillo
- Puente Negro
- num\_genes...
- num\_genes...
- max\_num\_ga...
- SDG
- functional...
- geometric...
- unintended\_h...

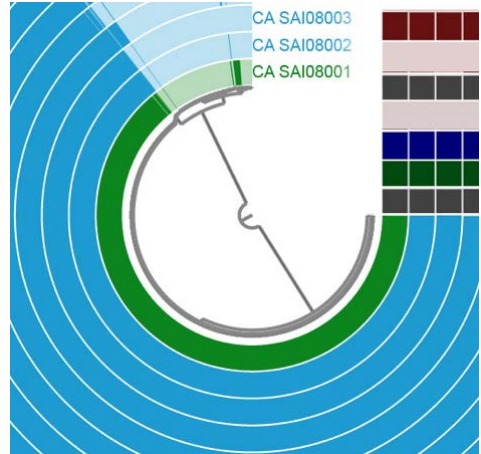
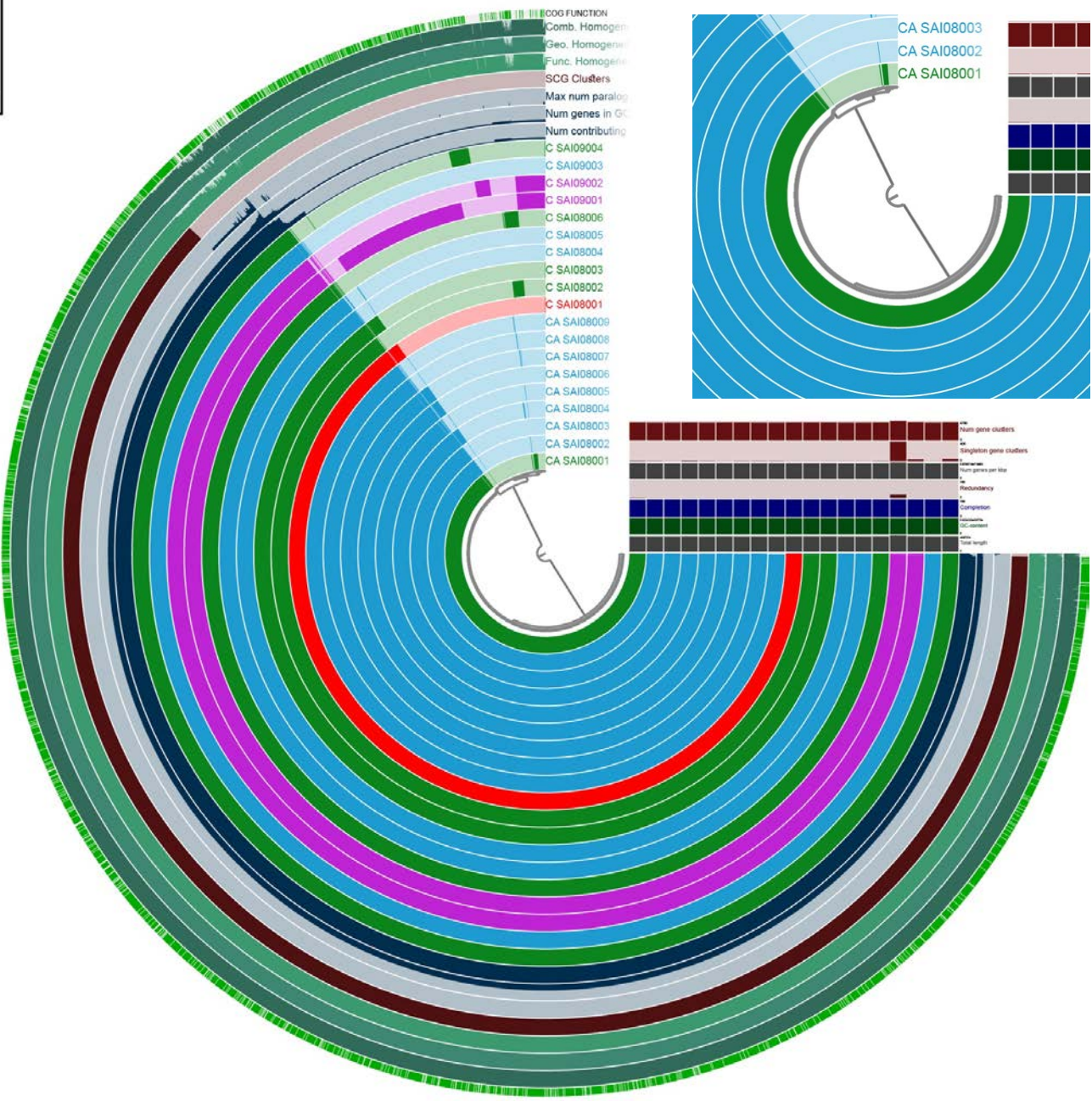




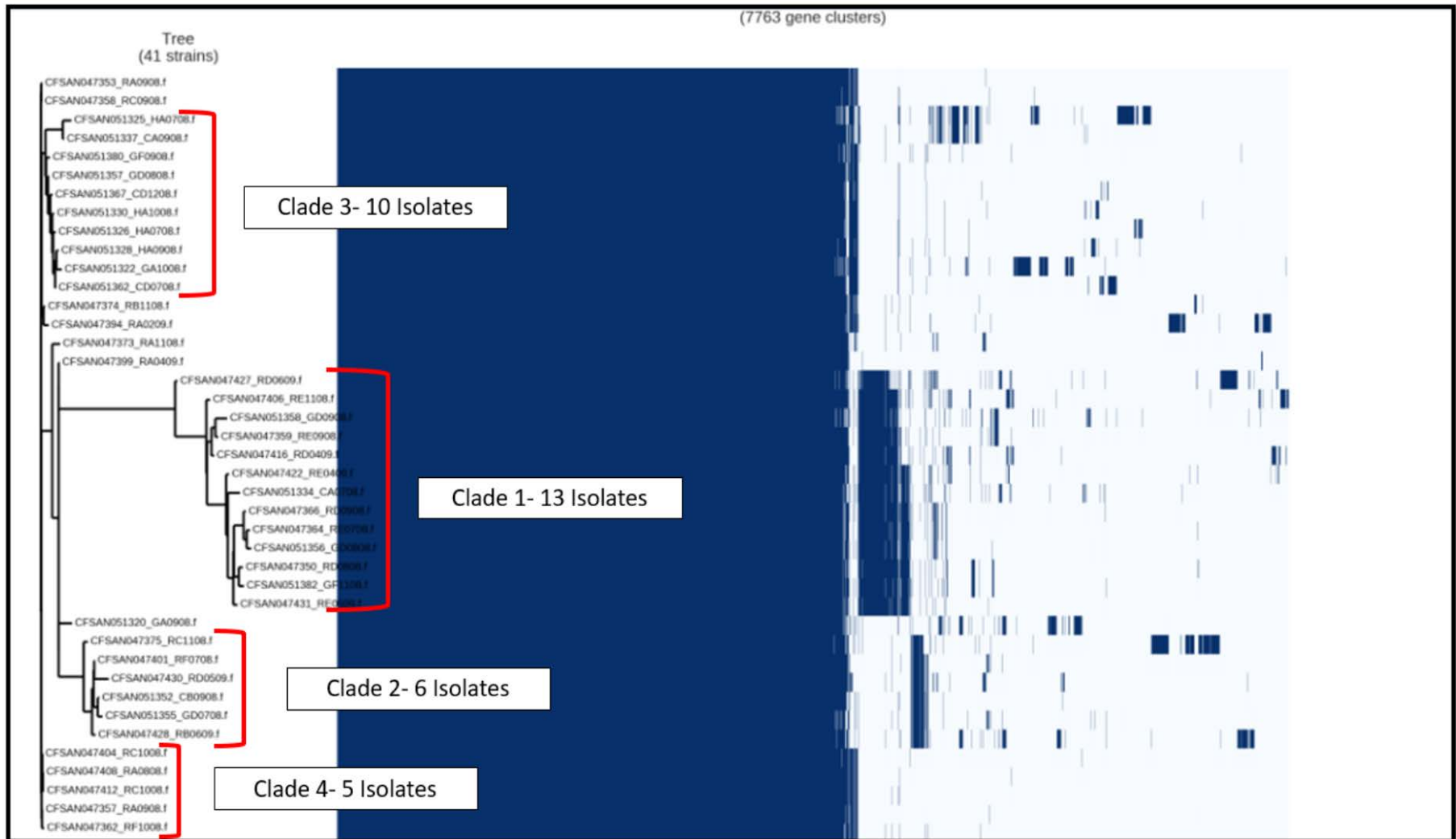
# SAINTPAUL PAN

Items order: Presence absence (D: Euclidean; L: Ward) | Current view: gene\_cluster\_presence\_absence | Samples order: custom

- num\_genes\_...
- max\_num\_pa...
- SCG
- functional...
- geometric\_...
- combined\_h...
- COG\_FUNCTI...



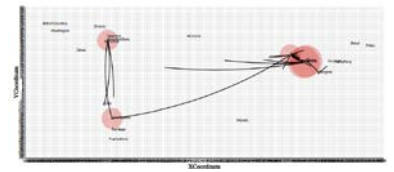
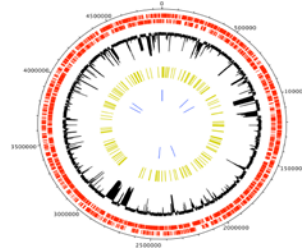
# Identification of Potential Niche-Specific Genes within the Accessory Genome of *Salmonella* Oranienburg



# Identification of Potential Niche-Specific Genes within the Accessory Genome of *Salmonella* Saintpaul

## Prophage occurrence within isolates of the three major clades

Isolate Name	Sample date	Sample location	Source	Clade #	Prophages (# of PHAST protein hits)					
					Gifsy-2 (47)	Gifsy-1 (27)	Fels-1 (23)	Burkho (17)	SEN4 (20)	SEN5 (20)
47351	26/08/08	El Castillo	River	A	1	1	1	0	0	0
47363	21/10/08	Jotagua	River	A	1	1	0	0	0	0
51344	12/08/08	Agua	Chicken	A	1	1	0	0	1	1
51346	09/09/08	Agua	Chicken	A	1	1	0	0	1	1
51347	23/09/08	Agua	Chicken	A	1	1	0	0	1	1
47396	25/02/09	Puente	River	B	1	1	1	1	0	0
47397	03/10/09	Puente	River	B	1	1	0	1	0	0
47361	10/07/08	Jotagua	River	C	0	0	0	1	0	0
47402	26/08/08	Jotagua	River	C	0	0	0	1	0	0
51336	26/08/08	Jotagua	Cow	C	0	0	0	1	0	0

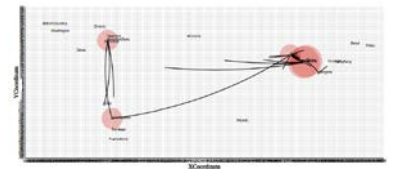
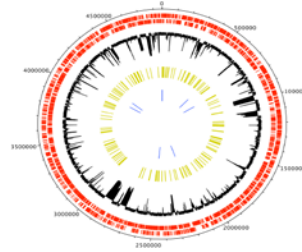


## CONCLUSIONS

According to results from MLST, isolates of *S. Oranienburg* and *S. Saintpaul* from Sinaloa belongs to a **single STs** which also included isolates from other countries.

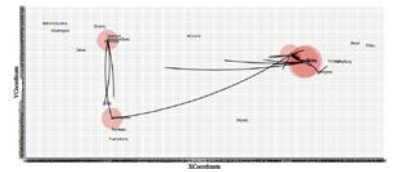
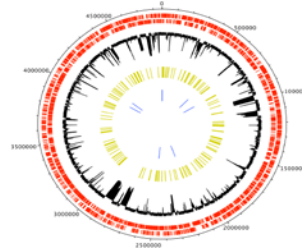
A deeper analysis onto the phylogenetic relationships within each ST revealed a more complex structure with clear differences between both serovars:

- ***S. Saintpaul*** isolates from Sinaloa were identified as member of a single cluster only comprised by Mexican isolates and **without connections** with isolates from anywhere in the world.
- ***S. Oranienburg*** populations from Sinaloa **shared position** in the phylogenetic tree with isolates from the USA.



## ANTIMICROBIAL RESISTANCE

- The diversity and prevalence of AMR genes detected within the genomes of serovars Saintpaul and Oranienburg analyzed in this study, indicated that *S. enterica* isolates from environmental sources **retain the potential for multi-drug resistance**.
- Due to the similarity of the antibiotic resistance profile found among *S. Oranienburg* and *S. Saintpaul* isolates, the AMR gene profiles have been shown to be of **limited applicability** for source tracking purposes.
- The detection of similar patterns of resistance in isolates from animals and water suggests that the **extensive use or misuse of antimicrobial agents**.



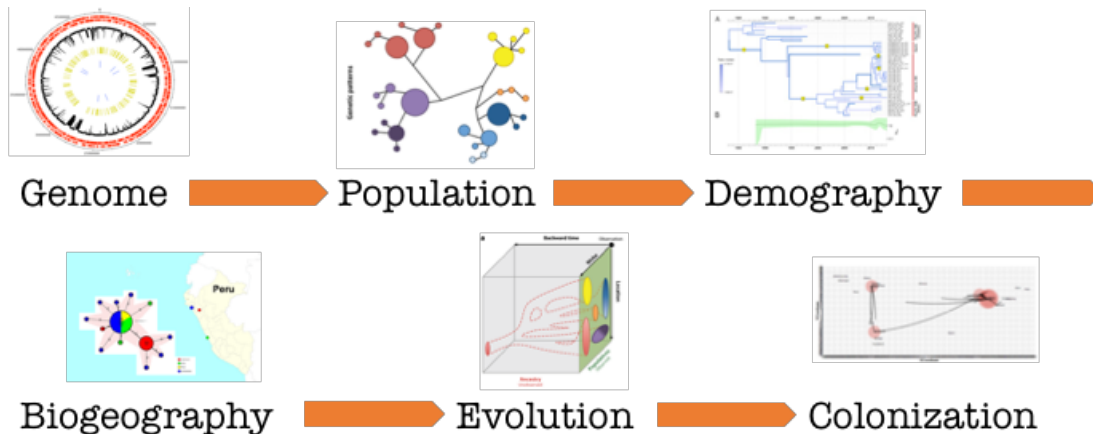
## SOURCE TRACKING

- Local populations of *S. Oranienburg* and *S. Saintpaul* in Sinaloa are composed ***by multi-resistant clonal groups moving frequently between asymptomatic domestic animals and non-host environments with no clear geographical barriers within sampling area*** investigated in this study.
- ***Domestic and farm animal activities near rivers are primary contributors*** to the persistence of specific clones found in environmental settings. The presence of strains with a high level of genetic similarity at different sampling sites revealed an ***effective circulation of Salmonella populations in the area potentially mediated by the movement of domestic animals and the dispersal through river flow.***



The application of a new generation of tools for microbial source tracking based on the use of genomic data identification of resistance and virulence genes, genomic characterization, pangenome construction,

... but would it be possible to reconstruct the biogeography of the dispersal.



RECONSTRUCTING THE LANDSCAPE OF SALMONELLA CONTAMINATION IN RIVERS USING A GENOMIC BAYESIAN FRAMEWORK



**Narjol Gonzalez-Escalona**  
**Maria Hoffmann**



**Joaquin Triñanes**



**Cristóbal Chaidez**  
**Juan Ramón Ibarra**  
**José Roberto Aguirre Sánchez**



UNIVERSITY OF  
**BATH**

**Michel Abanto**