Article | Open access | Published: 25 June 2024

# Spatial mapping of mobile genetic elements and their bacterial hosts in complex microbiomes

Benjamin Grodner, Hao Shi, Owen Farchione, Albert C. Vill, Ioannis Ntekas, Peter J. Diebold, David T. Wu, Chia-Yu Chen, David M. Kim, Warren R. Zipfel, Ilana L. Brito & Iwijn De Vlaminck

*Nature Microbiology* **9**, 2262–2277 (2024) | Cite this article

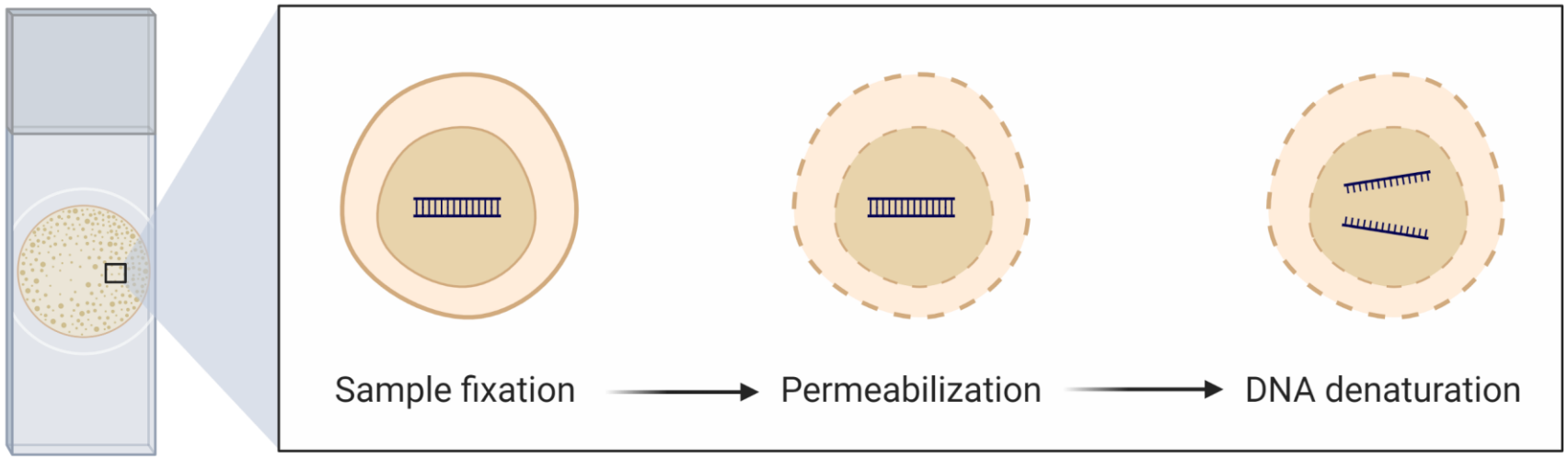


## Abstract

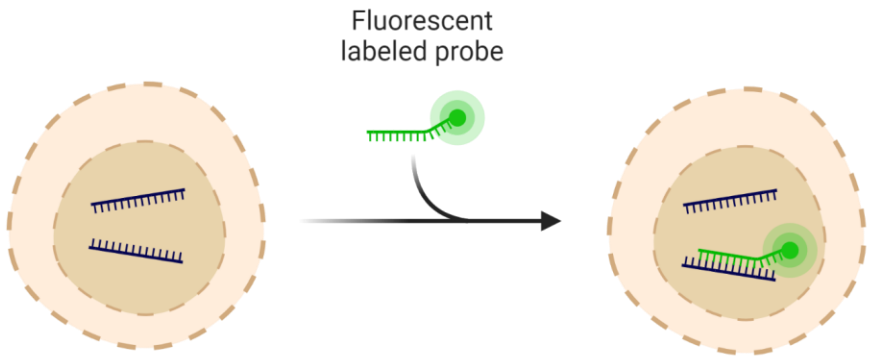
The exchange of mobile genetic elements (MGEs) facilitates the spread of functional traits including antimicrobial resistance within bacterial

The exchange of mobile genetic elements (MGEs) facilitates the spread of functional traits including antimicrobial resistance within bacterial communities. Tools to spatially map MGEs and identify their bacterial hosts in complex microbial communities are currently lacking, limiting our understanding of this process. Here we combined single-molecule DNA fluorescence in situ hybridization (FISH) with multiplexed ribosomal RNA-FISH to enable simultaneous visualization of both MGEs and bacterial taxa. We spatially mapped bacteriophage and antimicrobial resistance (AMR) plasmids and identified their host taxa in human oral biofilms.

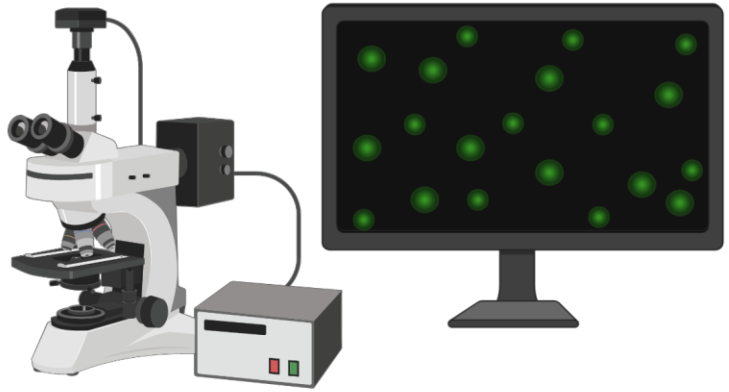
## 1 Sample preparation




## 2 Hybridization

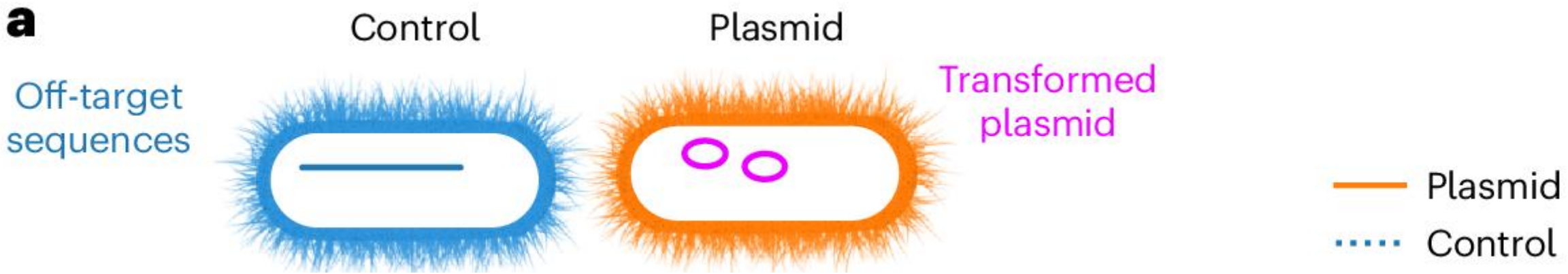



## 3 Imaging



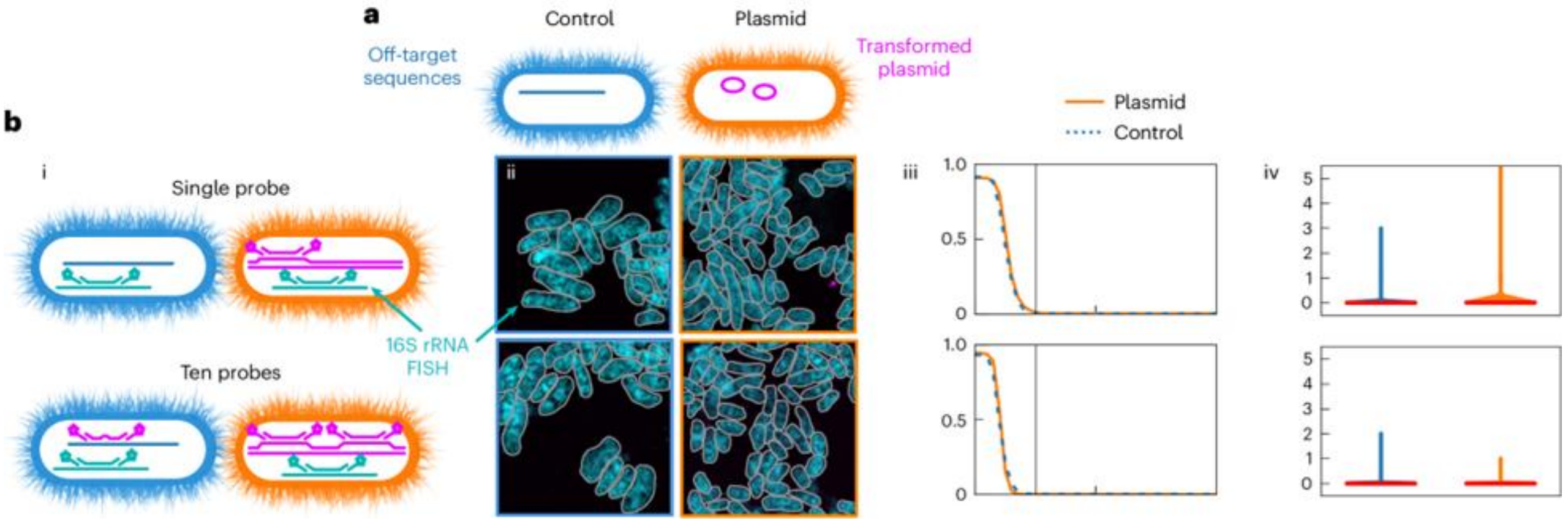
**single-molecule DNA fluorescence in situ hybridization (FISH)**
**Optimization of the method**

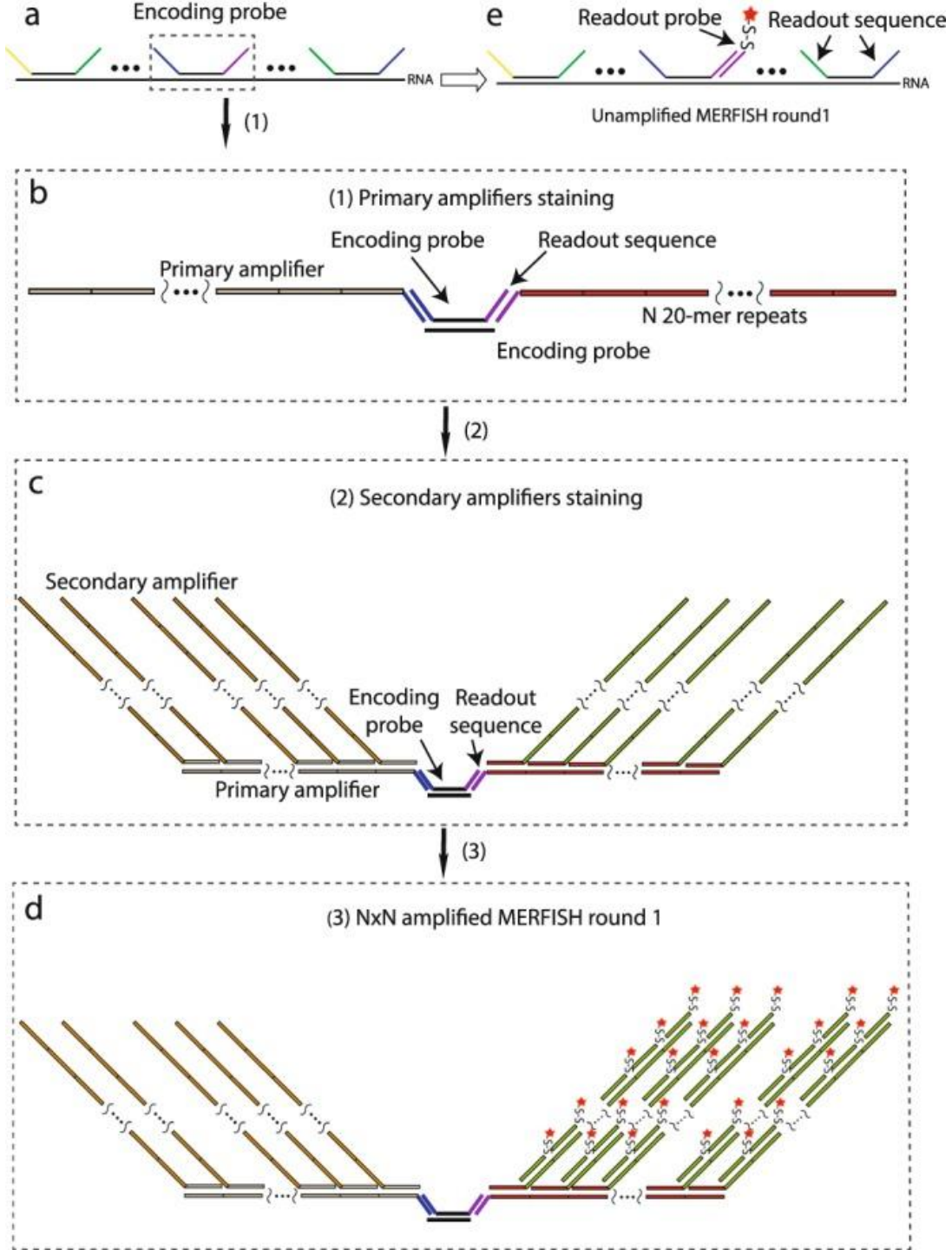
We used *Escherichia coli* transformed with pJKR-H-tetR plasmids encoding an inducible *GFP* gene as a model system to assess and optimize MGE-FISH on a confocal microscope (Fig. 1a)[10].
The chromosome was detected by 16S rDNA probe



Initial attempts using single and ten encoding probes yielded little to no separation between the signal in the plasmid and control samples (Fig. 1b, rows 1 and 2). This was expected given the photon noise and losses inherent in confocal microscopy as compared with a wide-field microscope[11,12].

We designed FISH probes for the non-coding strand of the *GFP* gene, used non-transformed *E. coli* as a negative control and tested sixdifferent FISH protocols. Initial attempts using single and ten encoding probes yielded little to no separation between the signal in the plasmid and control samples





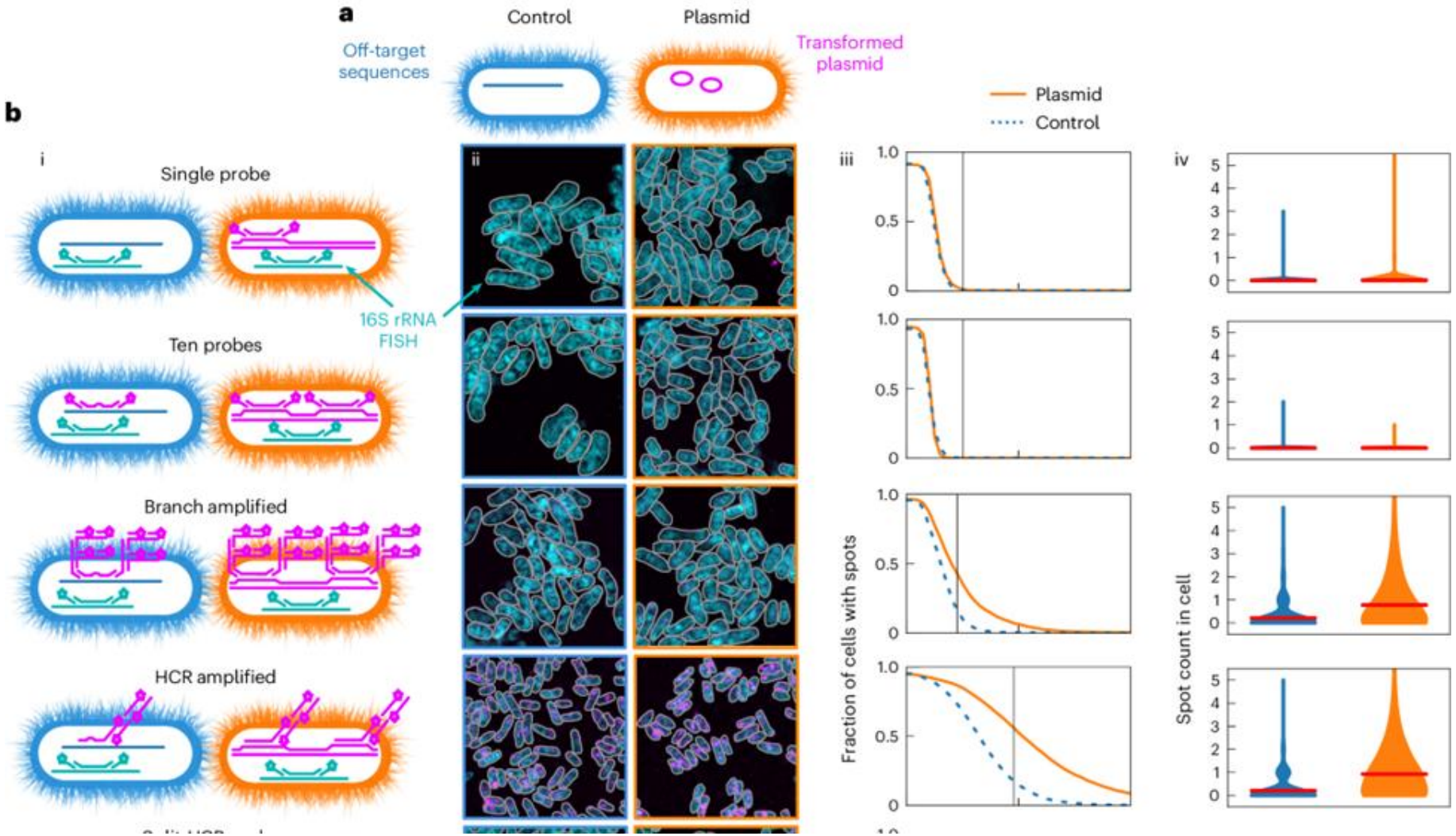

# a branched DNA (bDNA) amplification

Each encoding probe contains a 30-mer 'target region', whose sequence is complimentary to a region of the target RNA, and multiple 20-mer 'readout' sequences

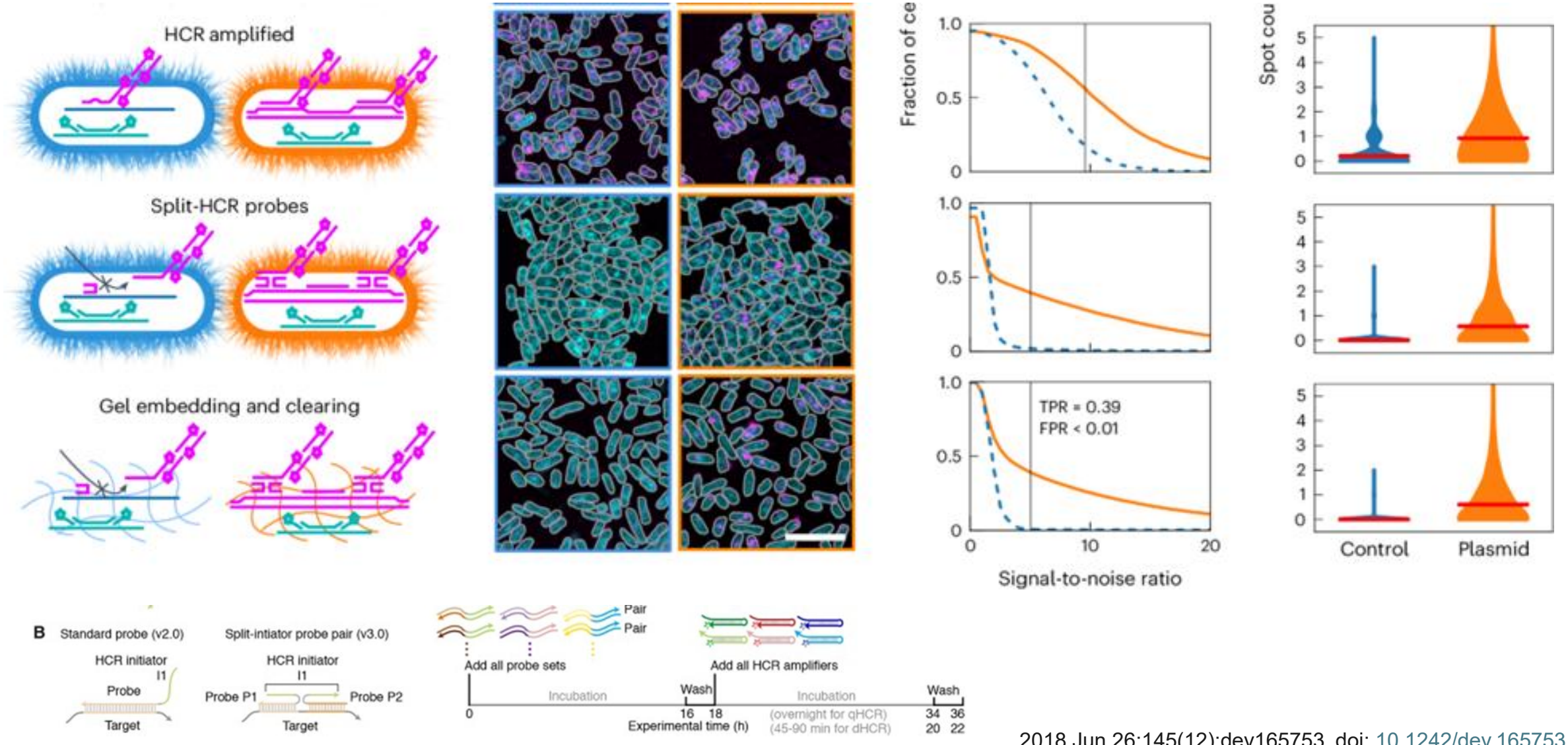
The primary amplifiers have a complimentary sequence to the readout sequence on encoding probes (blue or purple lines) and N 20-mer repeating sequences unique to each primary amplifier (tan or red lines). (**c**) Schematic depiction of the binding of secondary amplifiers to the primary amplifier's repeating regions. The secondary amplifiers have a complimentary sequence to one of the 20-mer repeating sequences on the primary amplifiers (tan or red lines) and N 20-mer repeating sequences unique to each secondary amplifier (orange or green lines).

Xia C, Babcock HP, Moffitt JR, Zhuang X. Multiplexed detection of RNA using MERFISH and branched DNA amplification. Sci Rep. 2019 May 22;9(1):7721. doi: 10.1038/s41598-019-43943-8. PMID: 31118500; PMCID: PMC6531529.

We next implemented two enzyme-free amplification methods to increase the signal[13,14]. Branched amplification yielded a higher true positive signal, albeit accompanied with a high background signal in the negative control (Fig. 1b, row 3). Hybridization chain reaction (HCR) similarly enhanced the signal at the expense of a high background in the control (Fig. 1b, row 4).

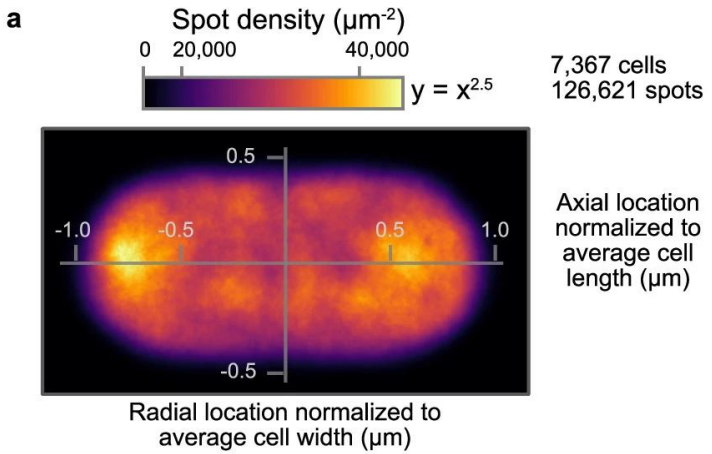


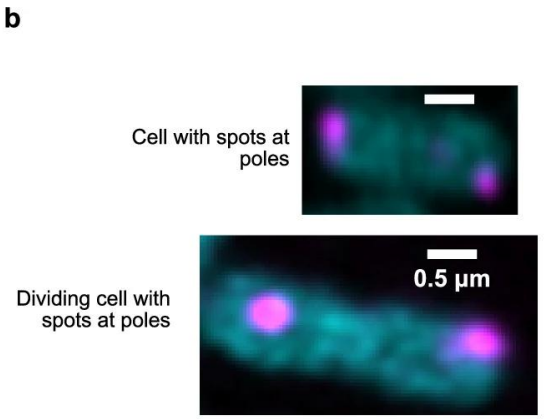
To improve specificity, we adopted a ‘split’ HCR method and used heat-denatured DNA and non-fluorescent ‘helper probes’ to stabilize the DNA[15,16]. This resulted in a significant reduction of the signal in the negative control (Fig. 1b, row 5). Last, to address autofluorescence in oral biofilms (as detailed below), we applied a gel embedding and clearing technique, in which nucleic acids in the sample are covalently anchored to a polyacrylamide gel, followed by clearing of proteins and lipids[17,18].

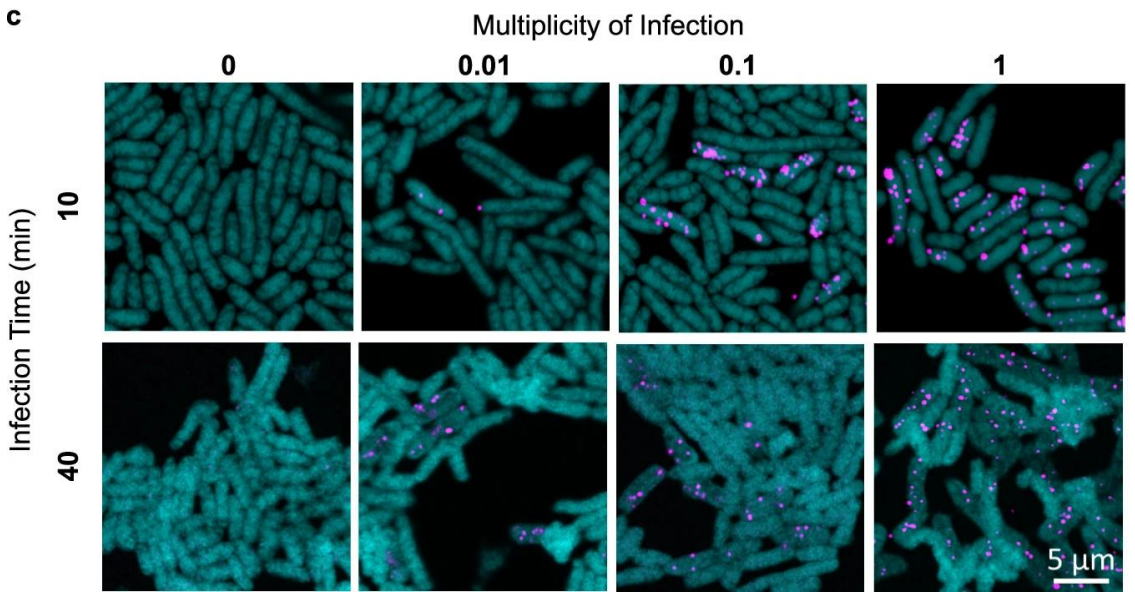


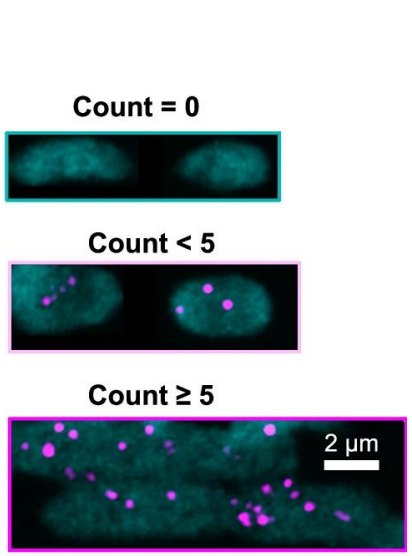

. 2018 Jun 26;145(12):dev165753. doi: 10.1242/dev.165753

We found that the plasmid density is ~50% higher on average at the poles compared with the centre (Extended Data Fig. 1a,b), in line with previous reports that plasmids have limited capacity to diffuse through the nucleoid at the cell centre and tend to cluster at cell poles[22,23].



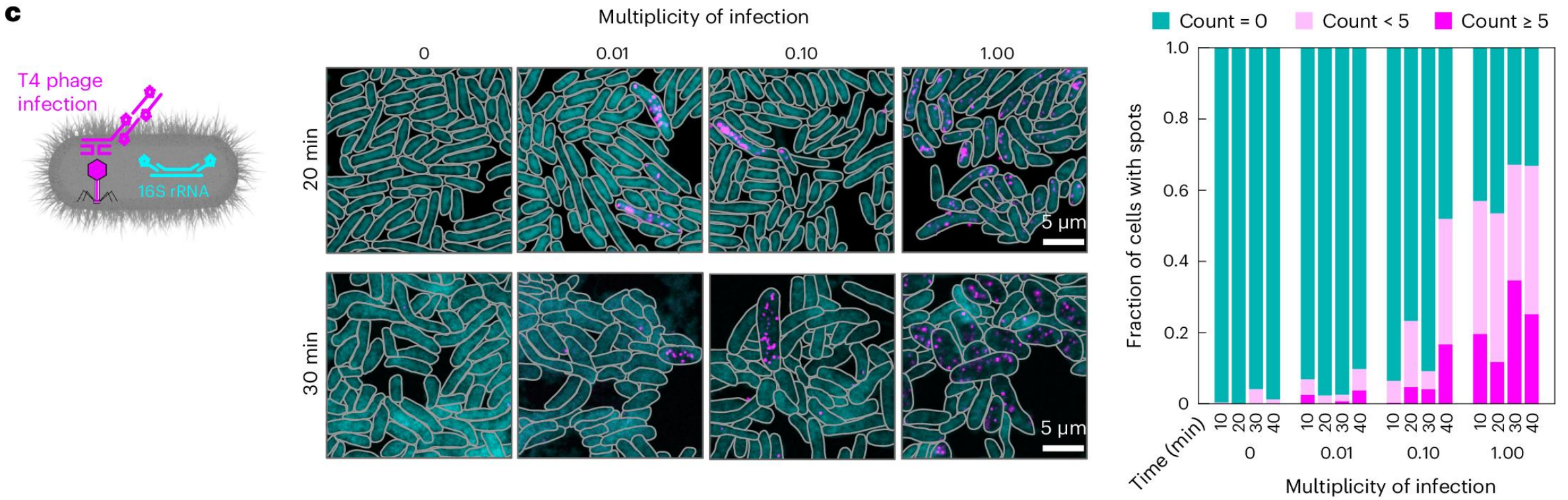



## Visualizing phage infection

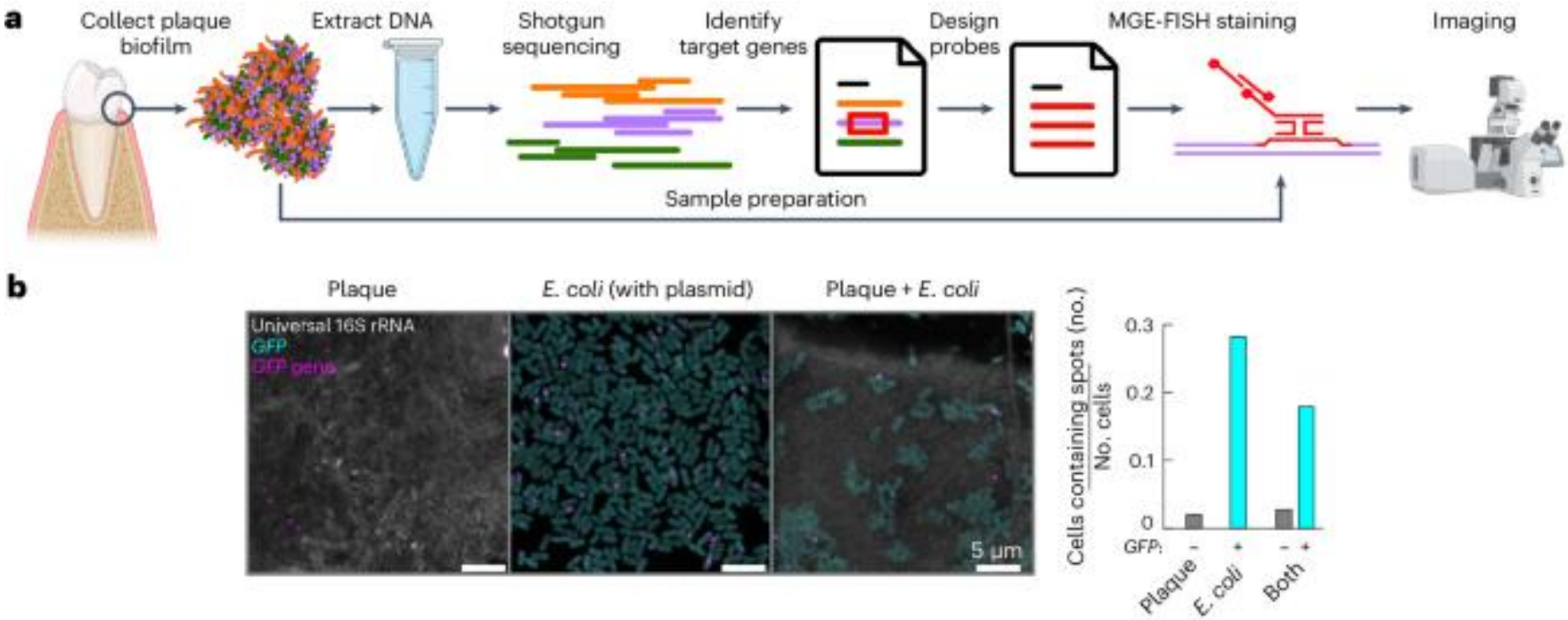
Building on the optimized MGE-FISH method (Fig. 1b, row 6), we turned our attention to visualizing T4 phage infection of *E. coli*. We staged infections at four multiplicities of infection (MOI 0, 0.01, 0.1 and 1) and fixed replicate cultures every 10 min over a 40-min period (Fig. 1c and Extended Data Fig. 1c). We designed FISH probes targeting the non-coding strand of the gp34 gene, which encodes a tail fibre protein, and quantified cells with 5 or more MGE spots, less than 5 spots and no spots (Fig. 1c and Extended Data Fig. 1d).



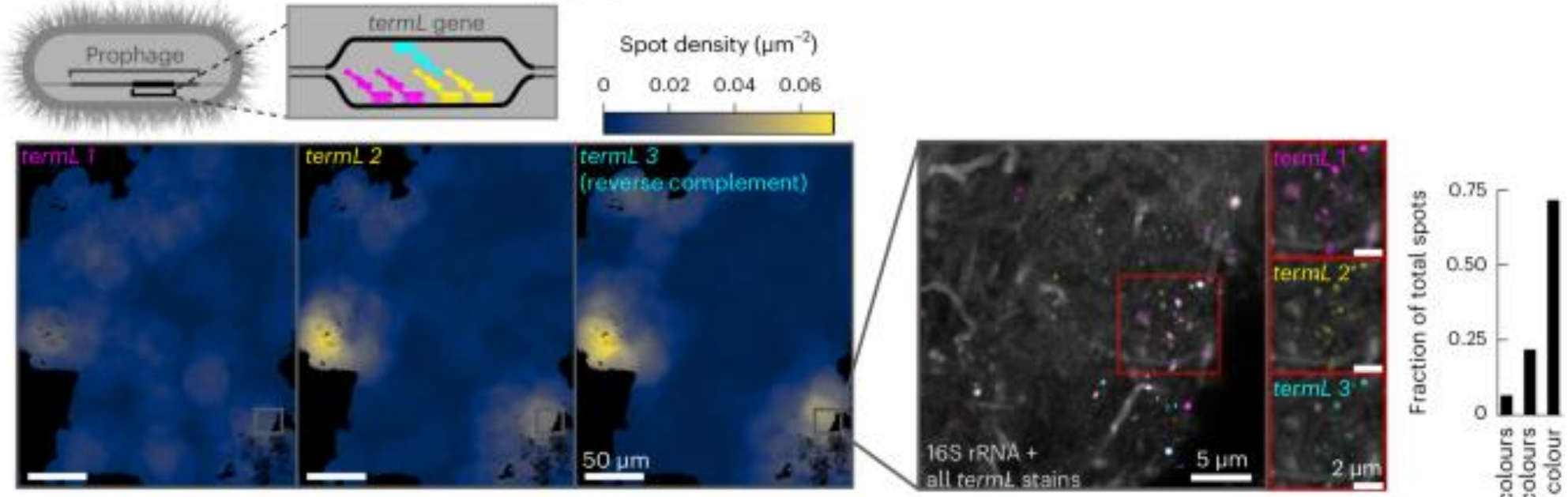

**c**, Left: diagram of MGE-FISH staining of *E. coli* infected by T4 phage. Middle: example images for four multiplicities of infection at 20 min and 30 min after introducing phage to the culture. Right: results of manual counting to classify cells into groups on the basis of the number of MGE-FISH spots.

## Mapping MGEs in oral plaque biofilms at high specificity

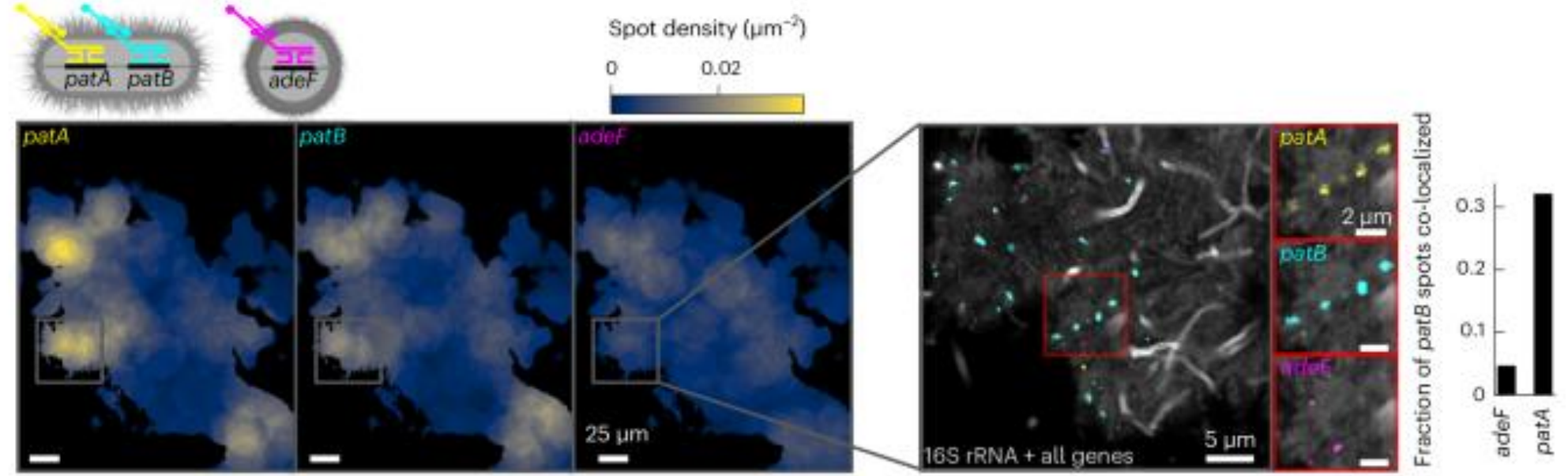
Next, we evaluated the ability of our MGE-FISH method to visualize the spatial distribution of MGEs in human oral plaque biofilms. To this end, we collected oral plaque biofilms from two healthy volunteers (A and B) and performed shotgun metagenomic sequencing on a portion of each sample, reserving the rest for imaging (Fig. 2a). As an initial controlled test of the method (Fig. 1b, row 6), we stained for the *GFP* gene in samples that contained mixtures of plaque and *GFP*-transformed *E. coli* (Fig. 2b) and demonstrated that the specificity remained high in plaque.



**d** *termL*: large terminase gene of *Caudoviricetes* prophage

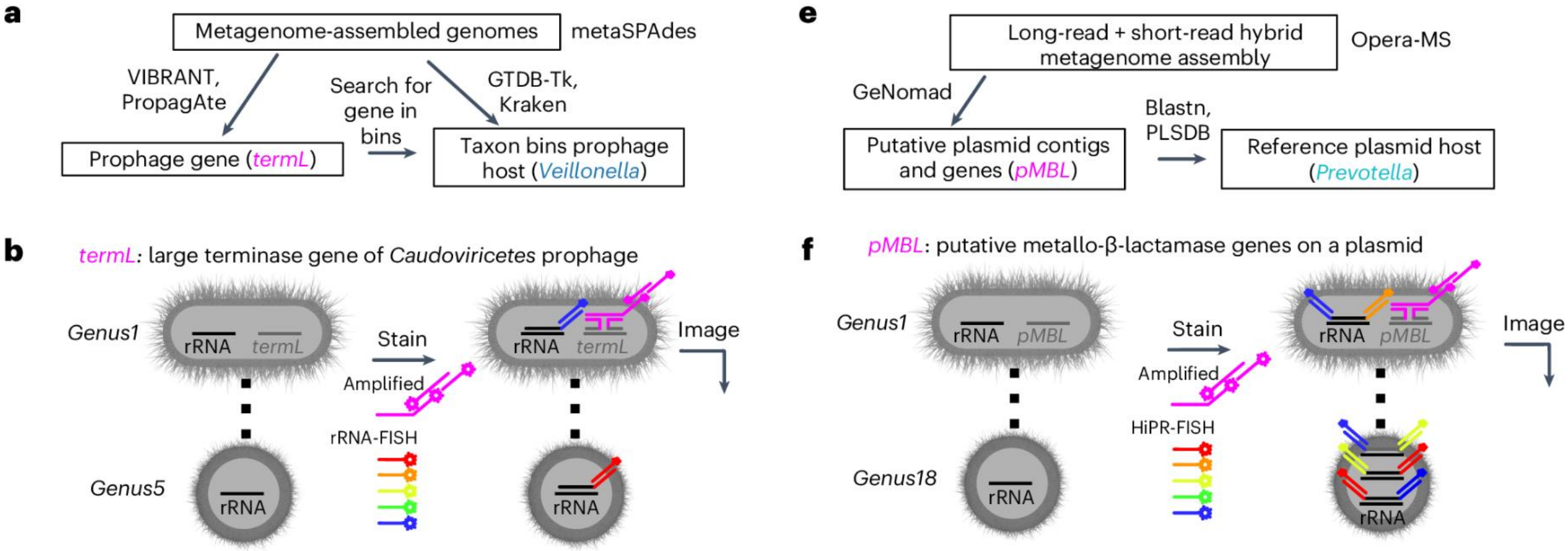



**e** *patA*, *patB*, *adeF*: antibiotic efflux pump subunits on chromosomes

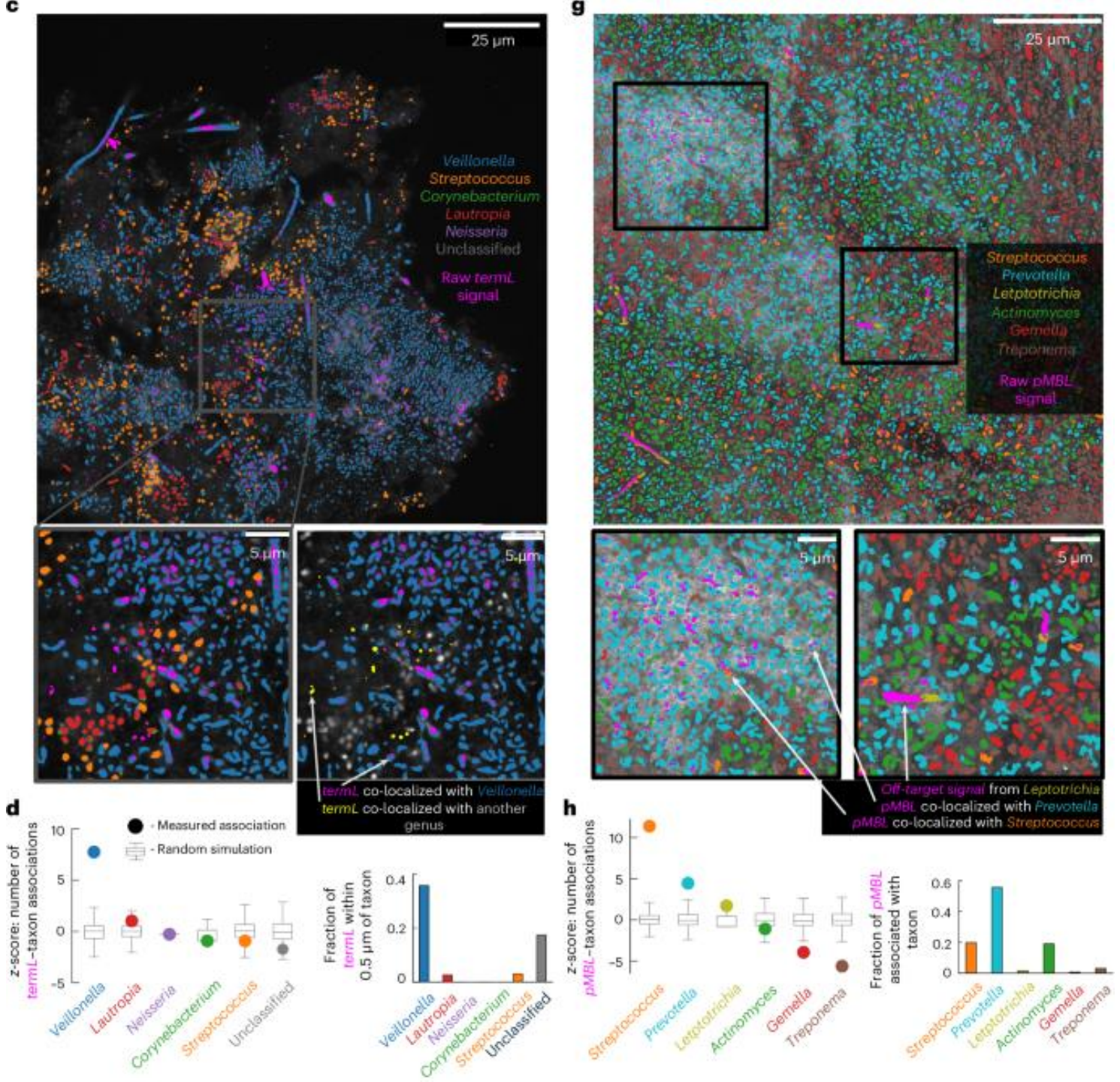


**Combined taxonomic mapping and MGE mapping**

We next overlaid MGE biofilm maps with taxonomic identity maps to associate MGEs with their host taxa. To start, we measured the taxonomic association of a highly abundant prophage of class *Caudoviricetes*, for which the metagenomic data and RefSeq alignment hinted at a strong taxonomic association with *Veillonella* (Fig. 3a ). We used rRNA-FISH to stain five common oral genera, *Veillonella*, *Streptococcus*, *Corynebacterium*, *Lautropia* and *Neisseria*, each with a different fluorophore, and we used MGE-FISH to stain the *termL* gene of the active prophage with a sixth fluorophore (Fig. 3b).



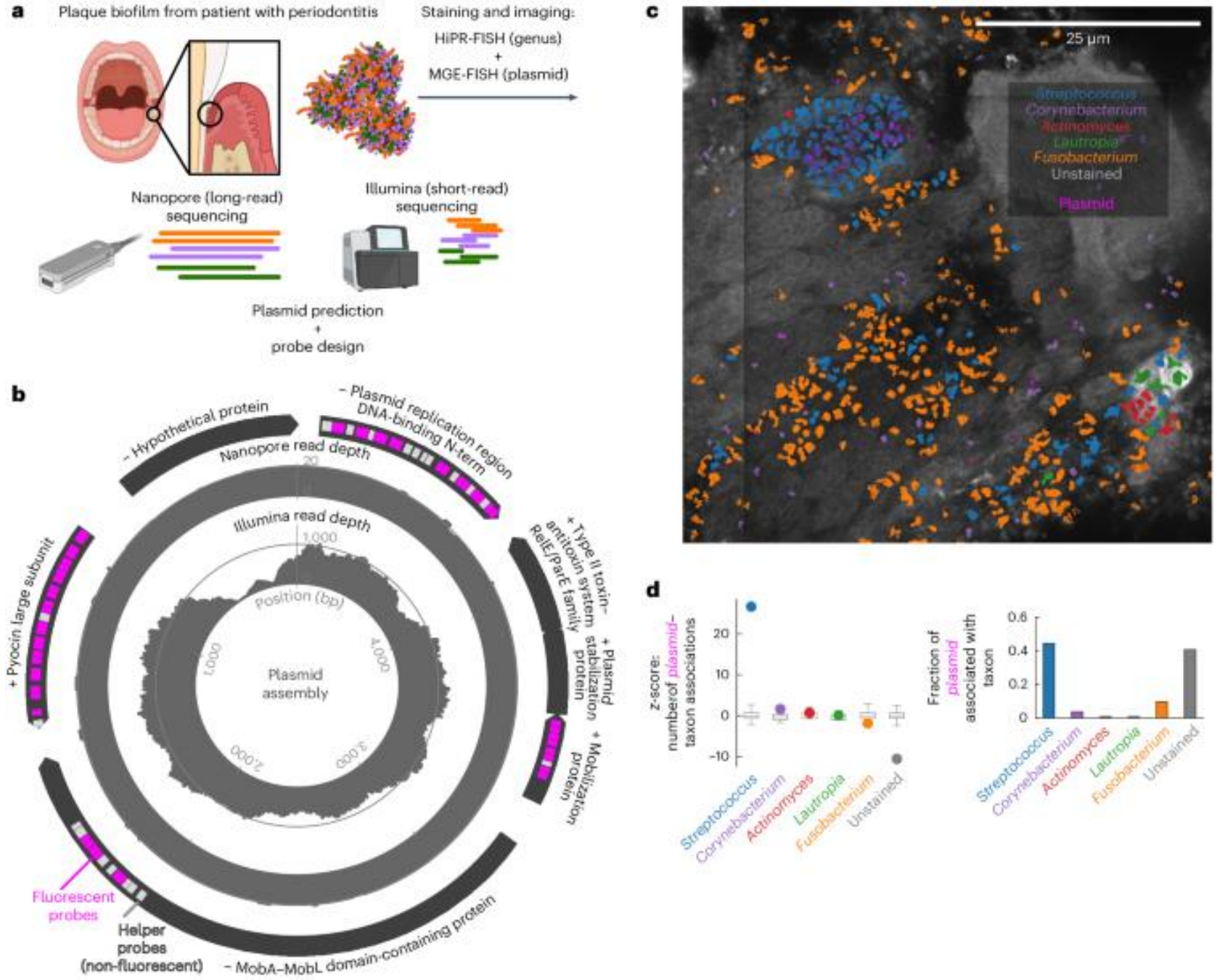

Next, we sought to confirm the host of a highly abundant plasmid discovered in the metagenomic data. We assembled contigs using combined long- and short-read sequencing and identified a highly abundant plasmid. Alignment of this contig to the plasmid database (PLSDB v.2023_11_23) showed that the plasmid had previously been observed in Prevotella nigrescens (Fig. 3e)36. We selected two genes from the contig that encode proteins with metallo-β-lactamase (MBL) domains as targets for MGE-FISH


c
25 µm
Veillonella
Streptococcus
Corynebacterium
Lautropia
Neisseria
Unclassified
Raw termL signal
5 µm
5 µm
termL co-localized with Veillonella
termL co-localized with another genus
g
25 µm
Streptococcus
Prevotella
Leptotrichia
Actinomyces
Gemella
Treponema
Raw pMBL signal
5 µm
5 µm
Off-target signal from Leptotrichia
pMBL co-localized with Prevotella
pMBL co-localized with Streptococcus
d
Measured association
Random simulation
z-score: number of termL–taxon associations
10
5
0
-5
Veillonella
Lautropia
Neisseria
Corynebacterium
Streptococcus
Unclassified
Fraction of termL within 0.5 µm of taxon
0.4
0.2
0
Veillonella
Lautropia
Neisseria
Corynebacterium
Streptococcus
Unclassified
h
z-score: number of pMBL–taxon associations
10
5
0
-5
Streptococcus
Prevotella
Leptotrichia
Actinomyces
Gemella
Treponema
Fraction of pMBL associated with taxon
0.6
0.4
0.2
0
Streptococcus
Prevotella
Leptotrichia
Actinomyces
Gemella
Treponema

Next, we investigated the taxonomic association of an unknown plasmid within a plaque biofilm of a patient diagnosed with stage 3 periodontitis. We combined long- and short-read sequencing to identify a complete plasmid with minimal homology to any sequence in the RefSeq database (v.220; Fig. 4a,b and Supplementary Table 10)[41].

The plasmid carried several predicted genes for mobilization and toxin–antitoxin systems.

However, we achieved complete coverage of the plasmid with the nanopore sequencing data, which allowed assembly of the full plasmid sequence (Fig. 4b). We designed MGE-FISH probes for the plasmid and combined this MGE-FISH stain with an 18-genera HiPR-FISH panel (Supplementary Table 12). We found that the plasmid was spatially associated with *Streptococcus* ($z$-score = 26.7, $P < 0.01$) and that *Streptococcus* formed small clusters within large patches of non-biofilm material (Fig. 4c,d).


a
Plaque biofilm from patient with periodontitis
Staining and imaging:
HiPR-FISH (genus)
+
MGE-FISH (plasmid)
Nanopore (long-read) sequencing
Illumina (short-read) sequencing
Plasmid prediction
+
probe design
b
– Hypothetical protein
– Plasmid replication region DNA-binding N-term
Nanopore read depth
20
Illumina read depth
1,000
position (bp)
1,000
2,000
3,000
4,000
Plasmid assembly
+ Type II toxin–antitoxin system RelE/ParE family
+ Plasmid stabilization protein
+ Mobilization protein
+ Pyocin large subunit
Fluorescent probes
Helper probes (non-fluorescent)
– MobA–MobL domain-containing protein
c
25 µm
Streptococcus
Corynebacterium
Actinomyces
Lautropia
Fusobacterium
Unstained
Plasmid
d
z-score: number of plasmid–taxon associations
20
10
0
−10
Streptococcus
Corynebacterium
Actinomyces
Lautropia
Fusobacterium
Unstained
Fraction of plasmid associated with taxon
0.6
0.4
0.2
0
Streptococcus
Corynebacterium
Actinomyces
Lautropia
Fusobacterium
Unstained

Using this method, we were able to make unique observations about MGE distributions across spatial scales in model bacteria and human oral plaque biofilms.

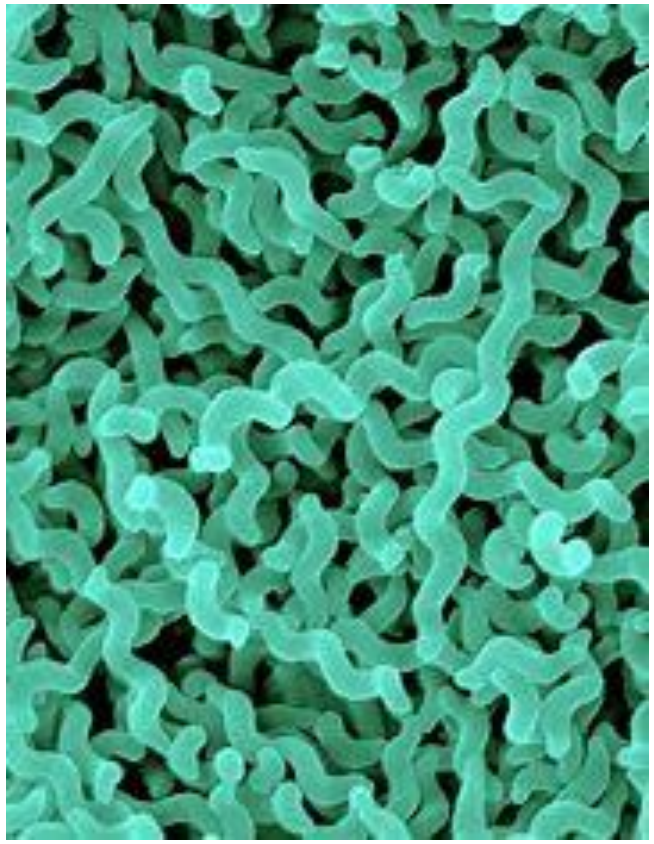
At the subcellular level, in vitro, we found that high-copy plasmids without partition systems show fewer puncta than expected and localize to the poles of the cells, which supports the idea that these plasmids bunch together within the cell and do not diffuse readily in the nucleoid.

We also showed that there are dramatic changes in cell shape and ribosome density associated with the number of copies of a replicating phage in *E. coli*, providing unexpected insight into the physical response of cells to infection. At the 10–100 µm scale in plaque biofilms, we demonstrated that AMR genes on plasmids and chromosomes can form clusters. We further observed clustering of two prophages at the same scale in plaque biofilms, with clusters of host cells isolated from each other by intervening non-host cells. Spatial clusters of prophages and AMR genes result from either short-range MGE exchange in dense clusters of host cells or clonal expansion of MGE host cells, but we cannot distinguish between these two possibilities with MGE-FISH.

We suggest that long-range (>100 µm) horizontal transfer of MGEs between clusters of host cells is limited by the need for MGEs to diffuse through the non-host biofilm.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 24 June 2021
doi: 10.3389/fmicb.2021.665858



# Multi-Omics Approach Reveals the Potential Core Vaccine Targets for the Emerging Foodborne Pathogen *Campylobacter jejuni*

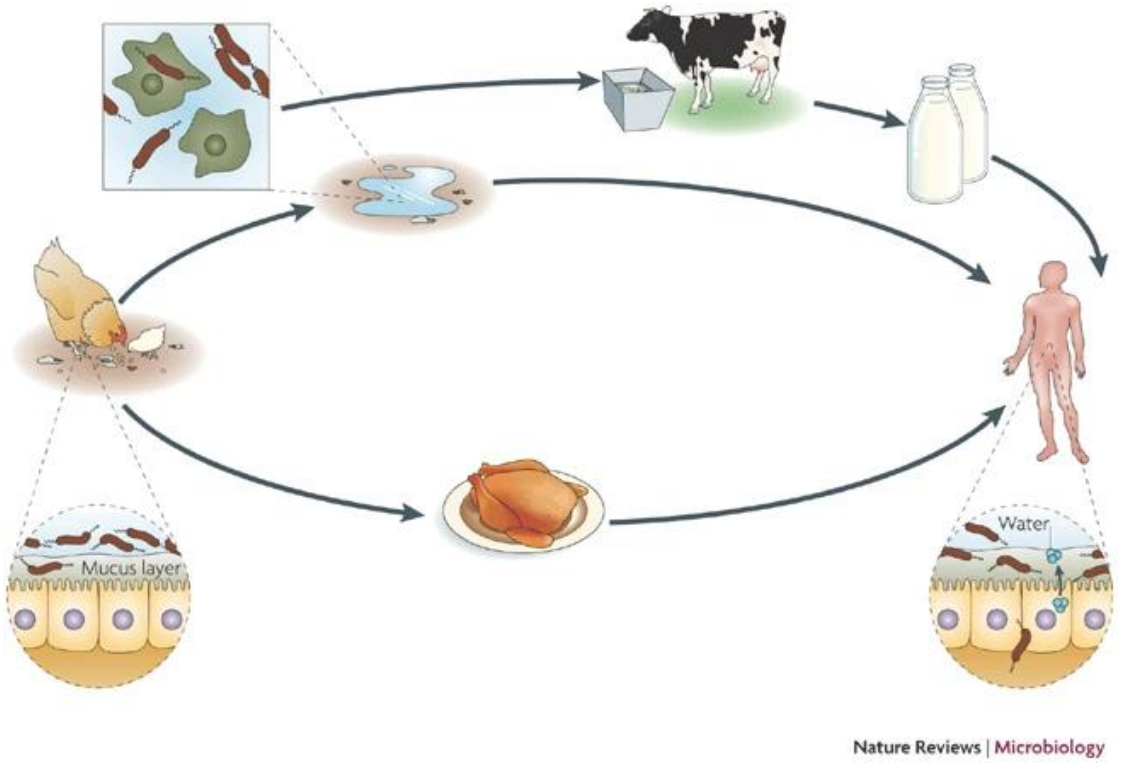
*Hengchun Cao[1], Hanxiao Xu[1], Chunhui Ning[1], Li Xiang[1], Qiufang Ren[1], Tiantian Zhang[1], Yusen Zhang[1]* and Rui Gao[2]*

*[1] School of Mathematics and Statistics, Shandong University, Weihai, China, [2] School of Control Science and Engineering, Shandong University, Jinan, China*

Thermotolerant campylobacters are the most frequent cause of bacterial infection of the lower intestine worldwide.

*C. jejuni* belongs to the epsilon class of proteobacteria, in the order *Campylobacteriales*; this order includes two other genera, *Helicobacter* and *Wolinella*. Like *C. jejuni*, members of these genera have small genomes (1.6–2.0 megabases)

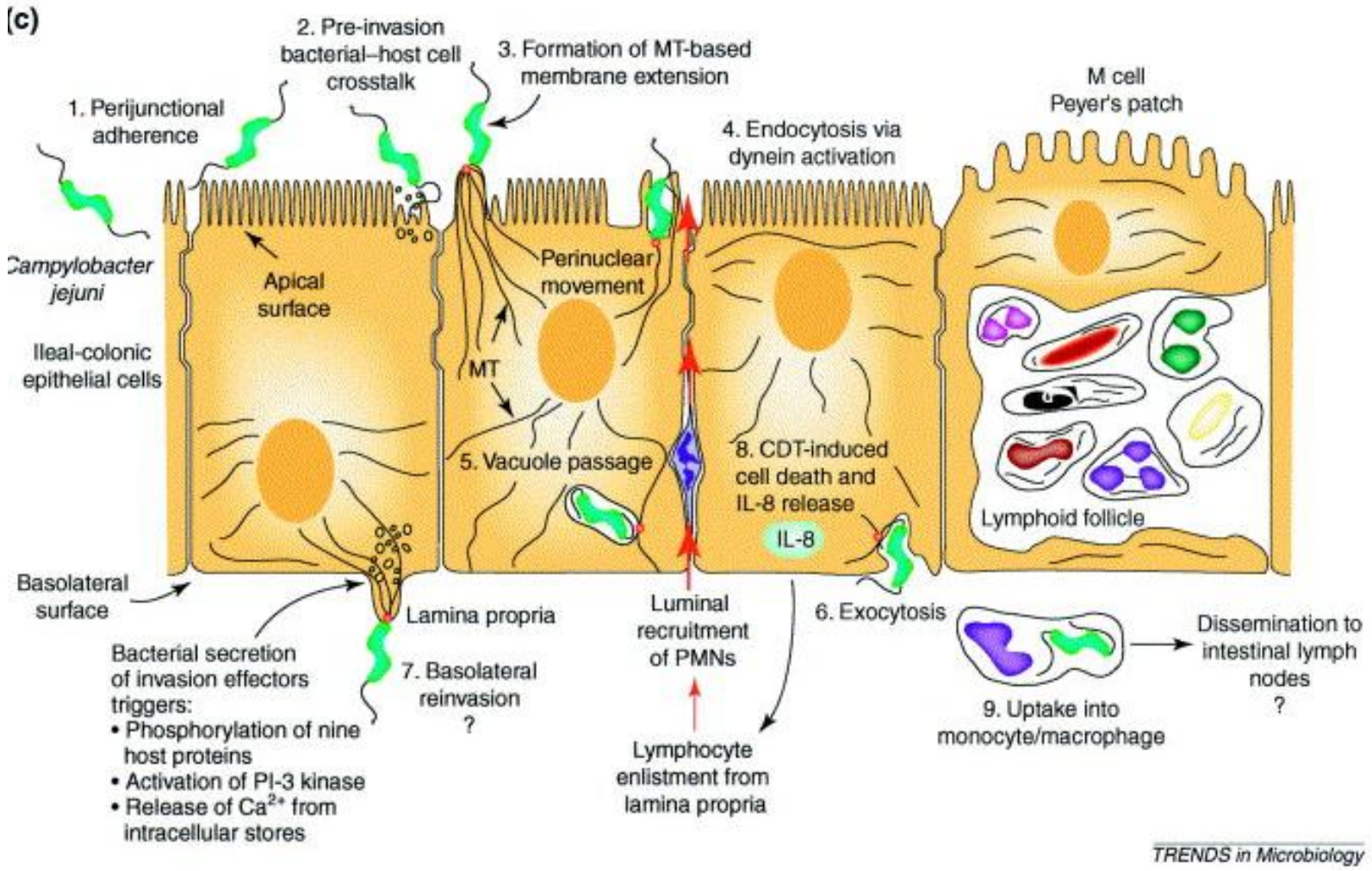
**Figure 1: The sources and outcomes of *Campylobacter jejuni* infection.**
Several environmental reservoirs can lead to human infection by *C. jejuni*. It colonizes the chicken gastrointestinal tract in high numbers, primarily in the mucosal layer, and is passed between chicks within a flock through the faecal–oral route. *C. jejuni* can enter the water supply, where it can associate with protozoans, such as freshwater amoebae, and possibly form biofilms. *C. jejuni* can infect humans directly through the drinking water or through the consumption of contaminated animal products, such as unpasteurized milk or meat, particularly poultry. In humans, *C. jejuni* can invade the intestinal epithelial layer, resulting in inflammation and diarrhoea.




Young KT, Davis LM, Dirita VJ. Campylobacter jejuni: molecular biology and pathogenesis. Nat Rev Microbiol. 2007 Sep;5(9):665-79. doi: 10.1038/nrmicro1718. PMID: 17703225.

# *Campylobacter jejuni* pathogenesis

The mechanism of pathogenesis comprises four main stages: **adhesion to intestinal cells, colonization of the digestive tract, invasion of targeted cells, and toxin production**.



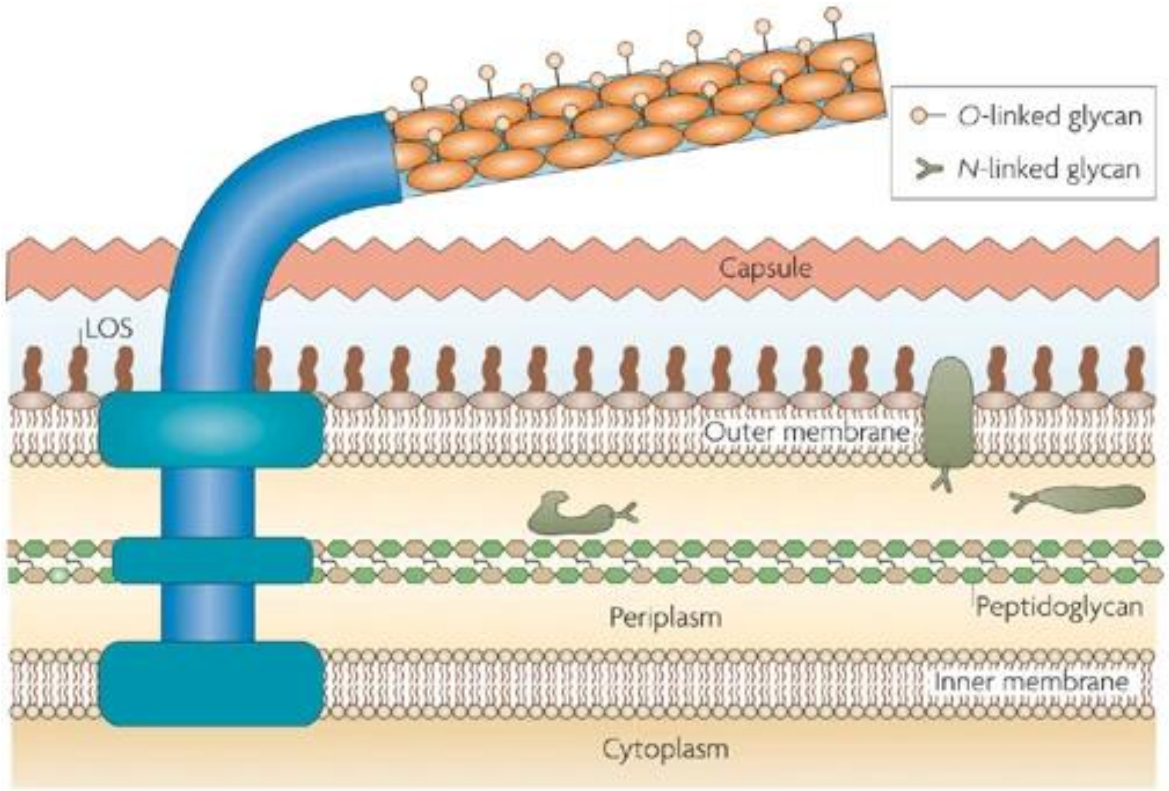
The determination of the complete genome sequence of several *C. jejuni* strains and plasmids has heralded the beginning of a new era of *C. jejuni* research. These projects have revealed the potential mechanisms by which *C. jejuni* associates with the host; for example, the complete sequencing of **pVir**, a plasmid that is found in some isolates of *C. jejuni*, has led to the identification of a **type IV secretion system** that has been demonstrated to have a role **in cell invasion and pathogenicity** in ferrets.
The publication of the genome sequence has also enabled the development of multiple genetic and genomic tools for use in *C. jejuni*, including microarrays, transposons for efficient random mutagenesis, signature-tagged mutagenesis, new reporter constructs and vectors for constructing in-frame deletion mutants and chromosomal point mutations.

**Genetic variation and natural transformation**.

*C. jejuni* displays extensive genetic variation, which has arisen from intragenomic mechanisms as well as genetic exchange between strains. Sequencing the genome of *C. jejuni* has revealed: the presence of hypervariable sequences that consist of homopolymeric tracts lack of clear homologues of many *E. coli* DNA-repair genes.

**Most of the hypervariable sequences are in regions that encode proteins that are involved in the biosynthesis or modification of surface-accessible carbohydrate structures, such as the capsule, lipooligosaccharide (LOS) and flagellum**.

The flagellin is modified by *O*-linked glycosylation. This modification is required for flagellar assembly and is, therefore, important for motility, virulence and epithelial cell adherence and invasion. The *N*-linked-glycosylation system modifies some periplasmic and outer-membrane proteins.

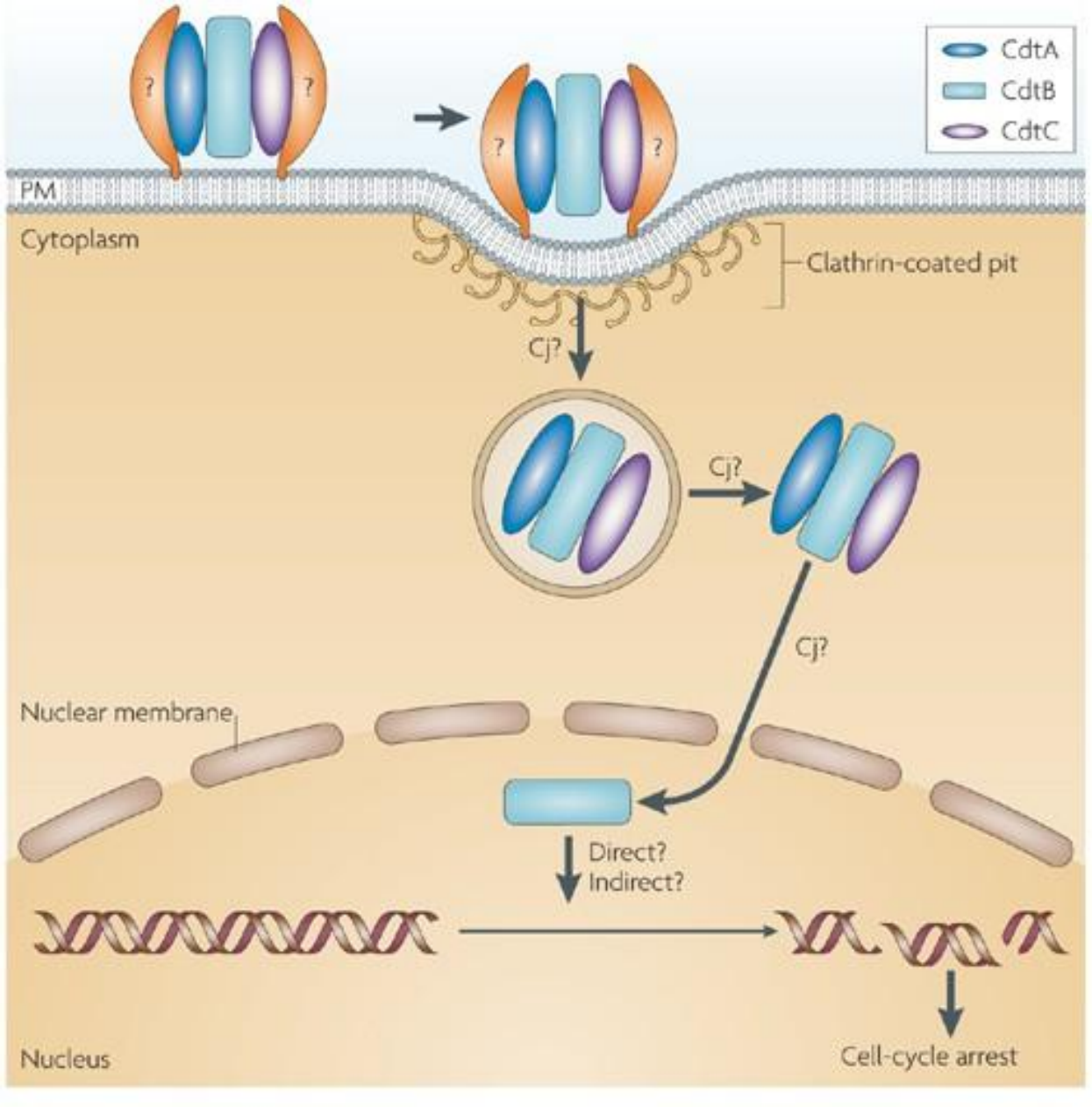


***Cytolethal distending toxin.***

*C. jejuni* produces cytolethal distending toxin (CDT), which is also produced by a diverse group of other bacterial species.

The toxin causes arrest at the $G_1$/S or $G_2$/M transition of the cell cycle, depending on the cell type.

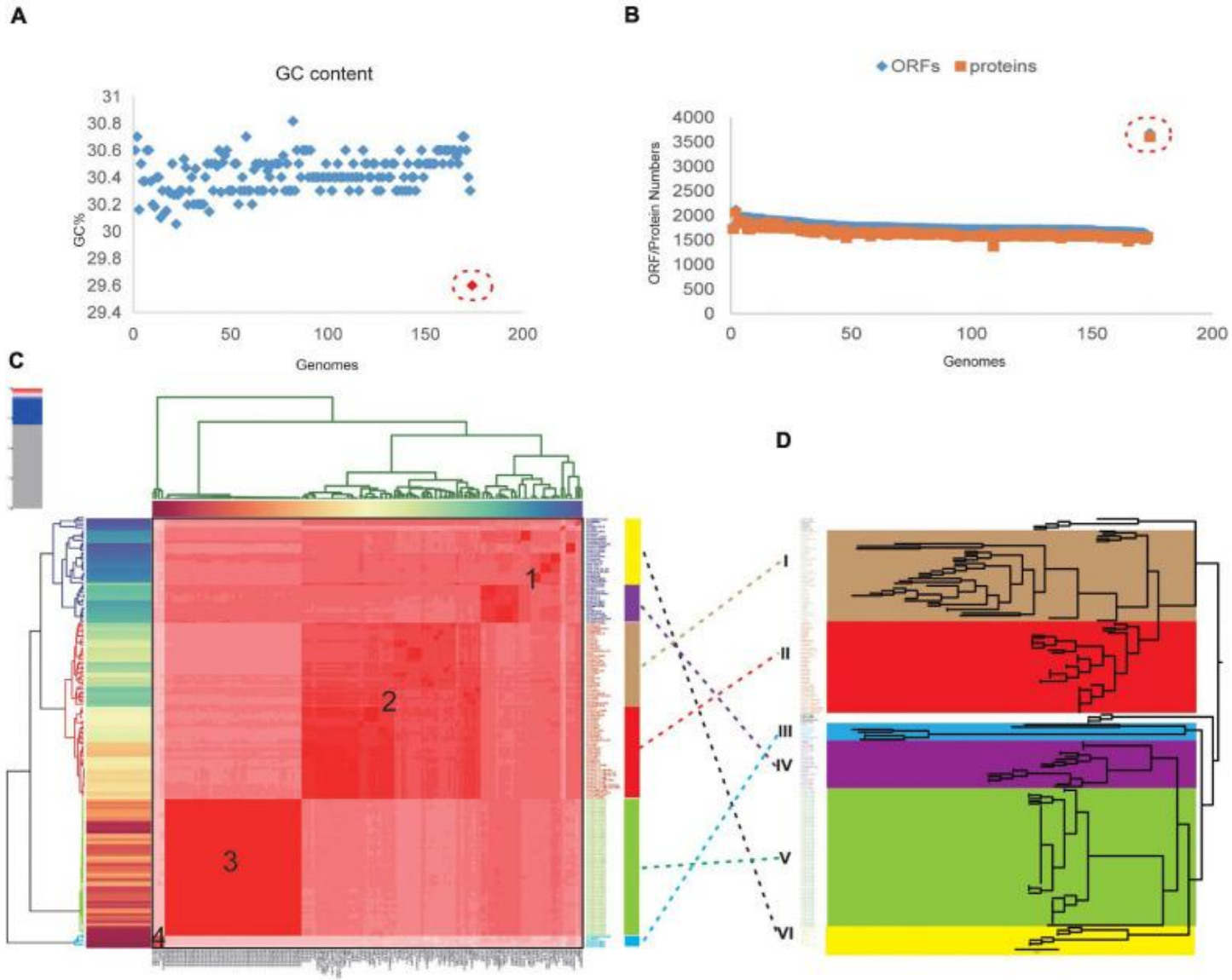
The active holotoxin is a tripartite complex of CdtA, CdtB and CdtC, although one study has indicated that CdtB and CdtC combined have some cytotoxicity without CdtA[1]



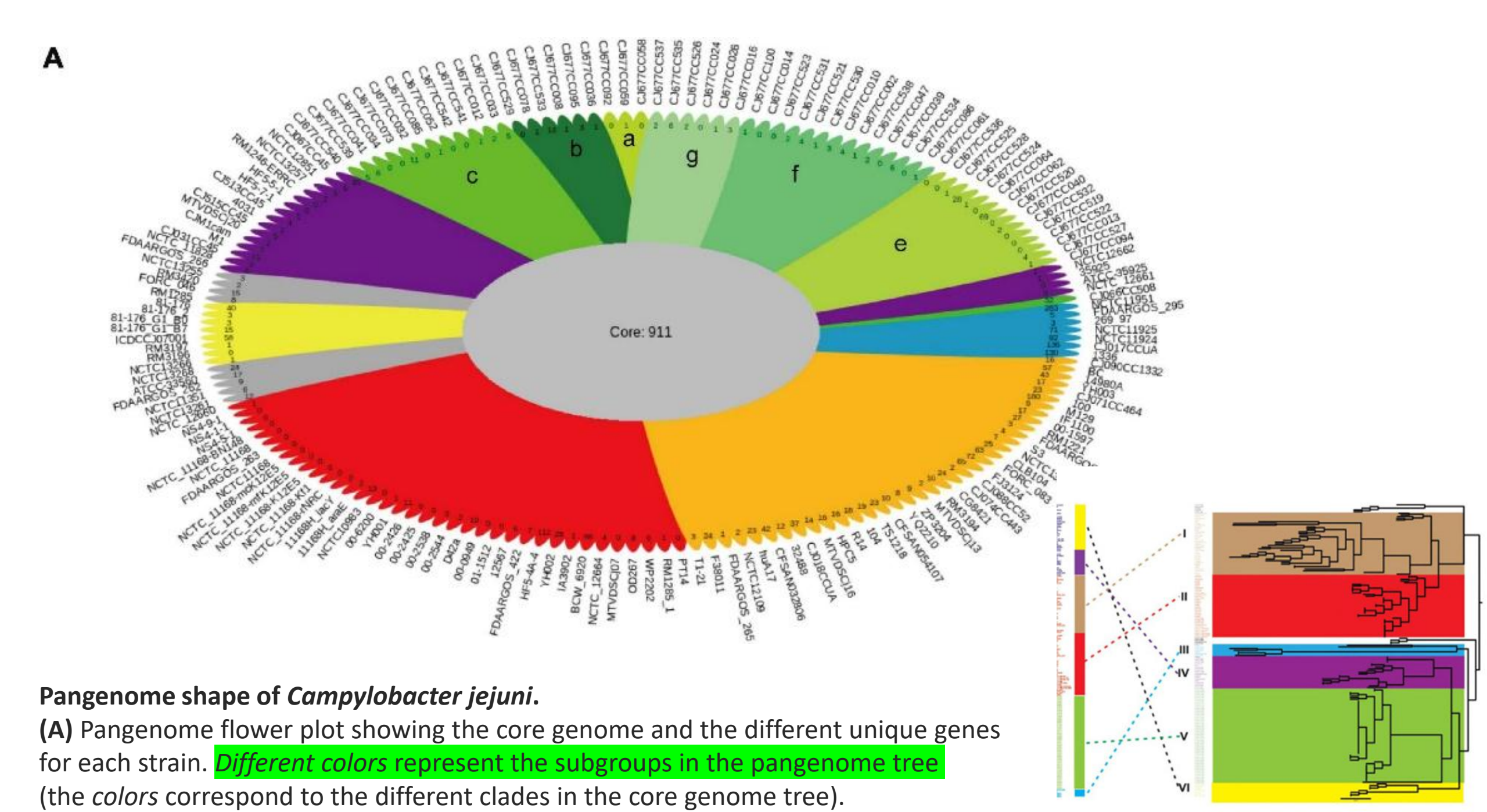
# *Campylobacter jejuni* genomics

A total of 174 complete genome sequences of *C. jejuni* strains collected from different geographic locations and isolation sources were preliminarily analyzed. To be consistent with the genomic data, all of the sequences were annotated using the software Prokka. The correct taxonomy classification is essential for obtaining high-quality pangenomes (Wu et al., 2020).

In order to determine the taxonomic status and obtain a high-quality pangenome of *C. jejuni*, the average nucleotide identity (ANI) values were firstly calculated to estimate the genetic relatedness among the strains. ANI has become one of the main genome options for DNA–DNA hybridization for taxonomic purposes. The previously suggested species threshold of 95% ANI can represent the same species. We found that the ANI value of the C. jejuni strain 414 is about 91%, which is obviously different from the other 173 strains and may be an incorrect classification.

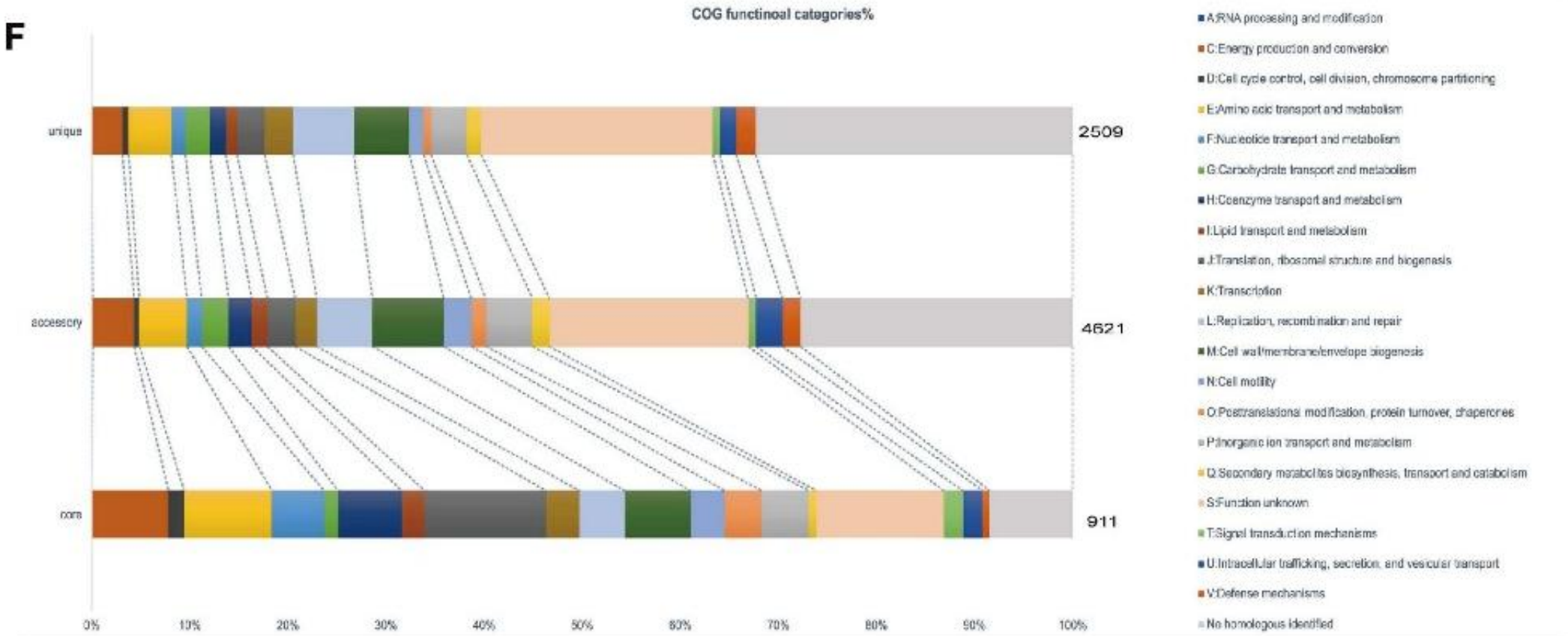



Using the whole-genome and core genome alignment concatenation approach, phylogenetic trees for the set of 173 genomes were constructed, the core genome tree could be divided into six main clades, in which nine strains were diverged independently of the other members.

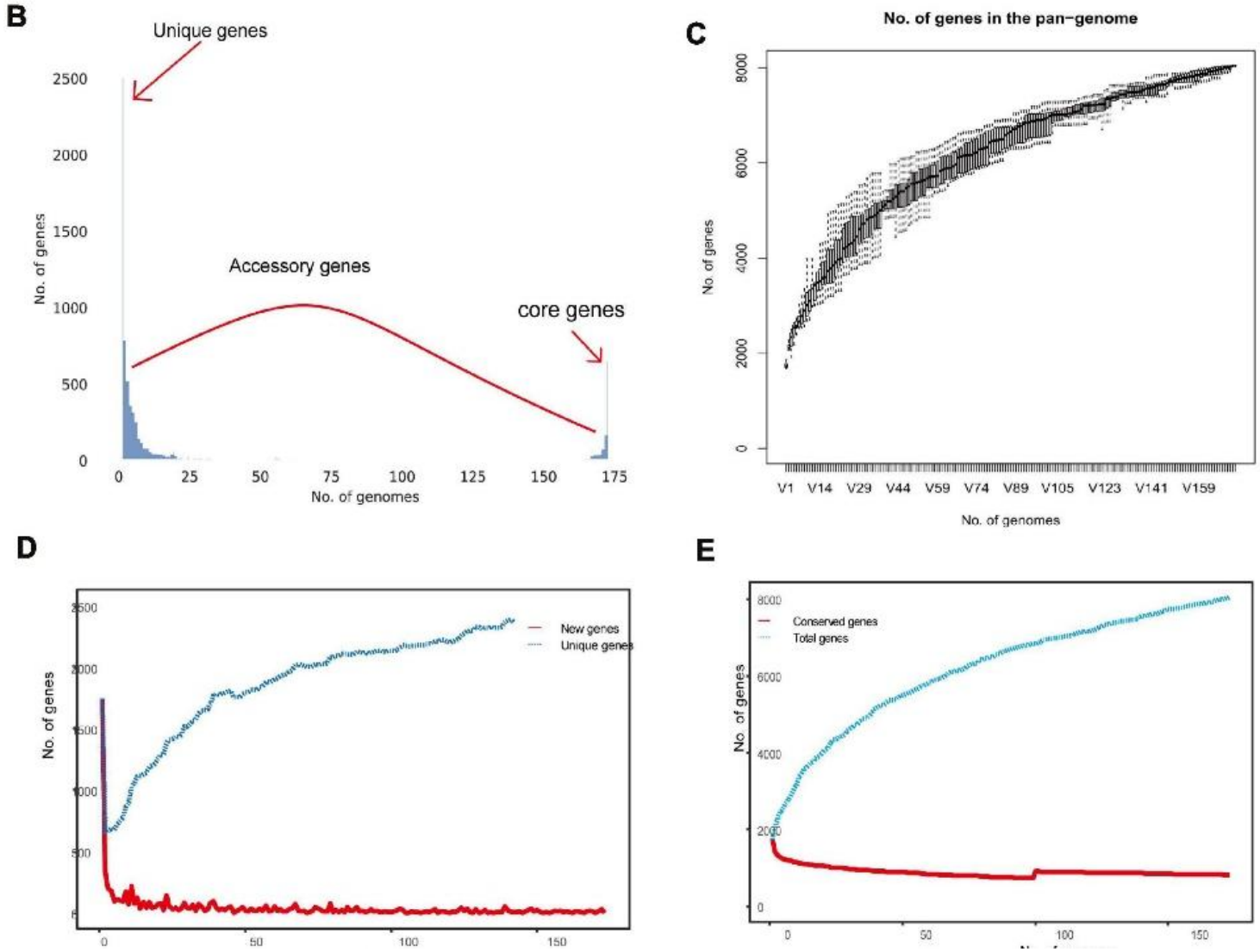



**Pangenome shape of *Campylobacter jejuni*.**

**(A)** Pangenome flower plot showing the core genome and the different unique genes for each strain. *Different colors* represent the subgroups in the pangenome tree (the *colors* correspond to the different clades in the core genome tree).




**Pangenome shape of *Campylobacter jejuni*.**
**(F)** Distributions of the Clusters of Orthologous Genes (COG) categories in the core, accessory, and unique genes without homologs were marked in *gray*.




open pangenome of 8,041 gene families was obtained with the correct taxonomy genomes.
The virulence property of the core genome was analyzed and 145 core virulence factor (VF) genes were obtained.

**Pangenome shape of *Campylobacter jejuni*.**
**(B)** Gene accumulation curves for the pangenome. **(C)** Histogram of the prevalence of the different gene families in the pangenome. A total of 8,041 non-redundant gene families identified in 173 genomes are based on their frequency distribution.


**Characterization of the core virulence factors (VFs).**

**(A)** Distributions of the virulence type categories of 145 VF proteins. The top three types are motility and export apparatus (*black*), colonization and immune evasion (*light blue*), and iron uptake (*brown*).

**(B)** Distributions of the Clusters of Orthologous Genes (COG) types of core VFs, with the largest COGs mainly distributed in cell motility (*N*) and cell component (wall, membrane, and envelop) biogenesis (*M*), which accounted for more than 35% of the 145 core VFs.

**(C)** Simulated location distributions of some core VFs based on Gene Ontology (GO) functional analysis that are mainly involved in motility, biosynthesis, metabolism, and transportation.

## Core VF Estimation for Essentiality and Non-host Homologs

Essential genes are composed of the minimum set of genes required to support cell life and have greater therapeutic potentiality.

The identification of essential genes is a key step in designing therapeutic targets for bacterial infections.

Among the 145 core VFs, 94 (~65%) were predicted as essential genes. These genes are mainly involved in biological processes like ATP binding, DNA binding, and transferase and permease activities.

Afterward, the essential core VFs were aligned with the human proteome to confirm whether there is any similarity between them.

74 proteins showed hits below the threshold value and were considered as non-host homologous proteins.

**These non-host homologous proteins can be preferably used for *C. jejuni* vaccine development to avoid autoimmune response or recombination and integration events in humans.**

Besides, proteins located in the periplasmic region, in outer membranes, and extracellularly are considered as effective vaccine candidates.

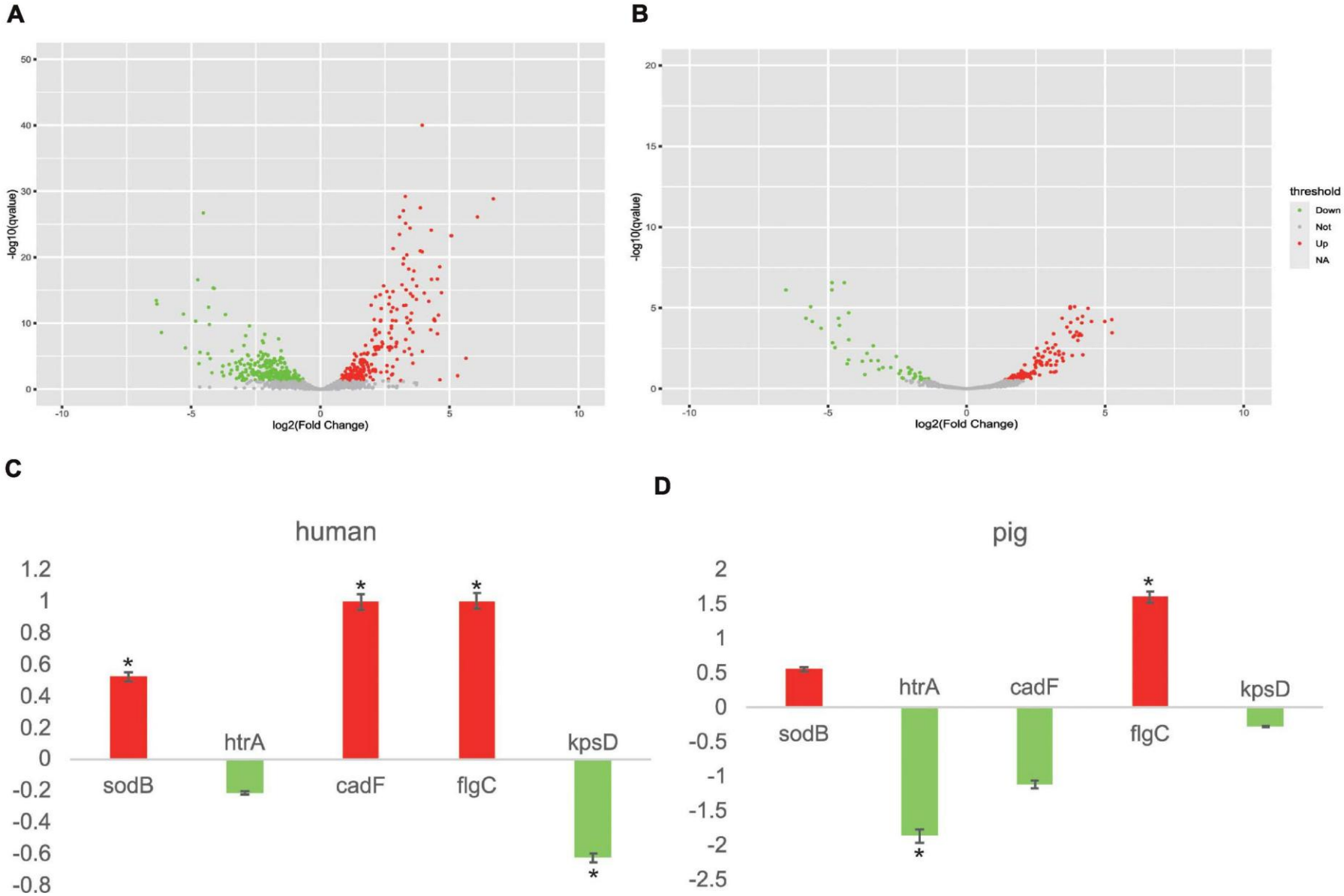
**The core VF proteins for subcellular location revealed that 47 proteins were cytoplasmic, 19 were located in the cytoplasmic membrane, one was in the outer membrane, three were unknown, and four were periplasmic**

It is known that outer membrane vesicles (OMVs) are a molecular complex consisting of lipopolysaccharides (LPS), outer membrane proteins, periplasmic proteins, lipids, and even cytoplasmic proteins, which are important vehicles for the simultaneous delivery of many effector molecules to host cells. Exposed proteins are often attractive targets for vaccine design, but sometimes not all proteins must be exposed to the surface, including some periplasmic proteins present in OMV preparations, which may also elicit an immunogenic response. Due to the role of OMVs in intestinal adhesion and invasion, and in regulating the dynamic interaction between host and pathogens, OMVs have become potential vaccine targets for a variety of intestinal pathogens. Therefore, the bacterial cell surface and secreted proteins, usually located in the extracellular, periplasmic, and outer membranes, could be more effective as vaccine candidates or diagnostic targets

| Protein name | Location | PsortB score | TMHMM prediction | Molecular weight (kDa) | VaxiJen score | VaxiJen prediction |
|---|---|---|---|---|---|---|
| SodB | Periplasmic | 9.44 | Outside | 24.81 | 0.5003 | Probable antigen |
| FlgC | Periplasmic | 9.44 | Outside | 18.30 | 0.4831 | Probable antigen |
| HtrA | Periplasmic | 9.76 | Outside | 51.01 | 0.5379 | Probable antigen |
| KpsD | Periplasmic | 9.44 | Outside | 60.84 | 0.4261 | Probable antigen |
| CadF | Outer membrane | 10 | Outside | 36.00 | 0.8043 | Probable antigen |

***Campylobacter jejuni* transcriptome analysis in human INT 407 and Caco-2 cells and the pig intestinal loop.**

the expression levels of 126 genes, including the 25 core VFs (**which include *sodB*, *cadF*, and *flgC***) were increased and the expression levels of 148 genes (**including 13 core VFs**) were decreased under human immune stress.



Upon functional genomics and immunological analyses, five core VF proteins with high antigenicity were selected as potential core vaccine targets for humans.

For the five selected proteins, nearly all of them had an apparent differential to the stress in human and pig.

As well, the results in the pig ligated intestinal loop model showed that the expression levels of 33 core VFs, including *flgC*, have been increased and those of 23 core VFs, including *htrA*, have been decreased.

**The oxidative stress response genes and the iron acquisition genes, including the potential vaccine targets *htrA*, *sodB*, and other core VFs such as *chuA*, *chuB*, and *chuD*, were expected to be decreased due to the intestinal mucus in the intestinal loop of pig.**

**This study found that the increased core VFs were mainly associated with the motility- and flagellar-related genes in both human and pigs and that the decreased core VFs were mainly related to iron transport system proteins.**

**These results indicate that the flagellar genes are important VFs, which are essential for *C. jejuni* motility and the secretion of virulence proteins.** The differences in the gene expressions could be caused by the different transcriptional responses by different hosts or the need for a certain reaction time after infection.

**The candidate proteins found in this study may be efficient vaccine targets both in human and other animals. With the development of more animal models, these core VFs can provide abundant gene resources, which may be beneficial to the study of the virulence mechanisms of *C. jejuni***

nature communications

Article

https://doi.org/10.1038/s41467-023-43854-3

# Global emergence of a hypervirulent carbapenem-resistant *Escherichia coli* ST410 clone

Received: 10 May 2023

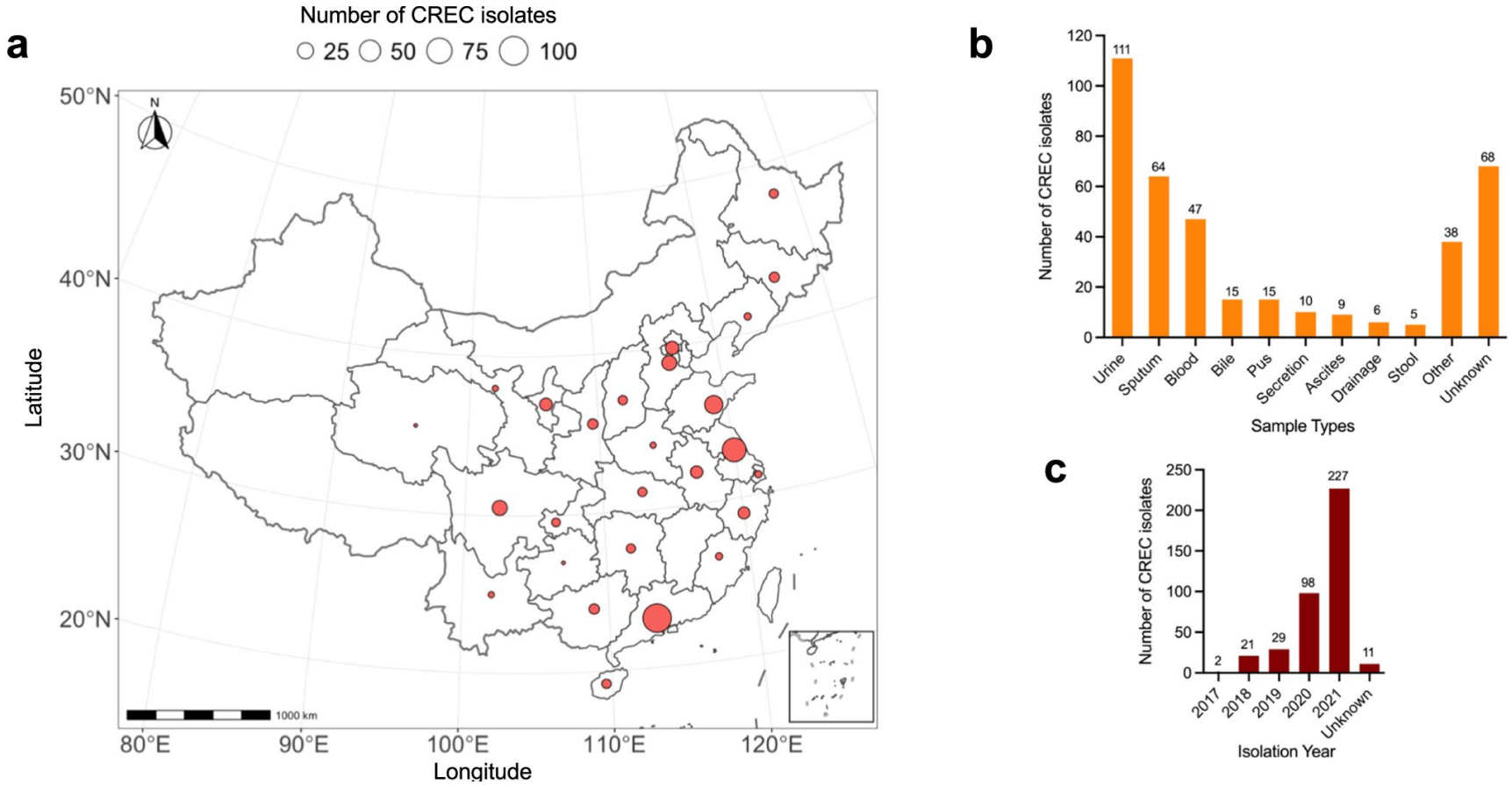
Accepted: 22 November 2023

Published online: 12 January 2024

Check for updates

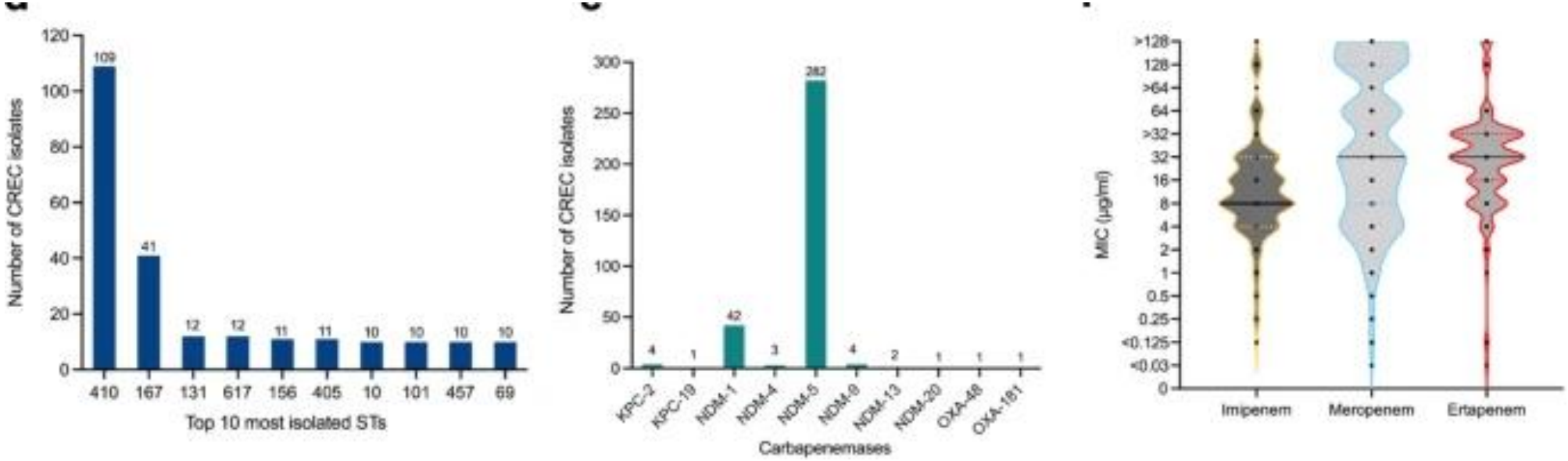
Xiaoliang Ba[1,8], Yingyi Guo[2,8], Robert A. Moran[3], Emma L. Doughty[3], Baomo Liu[4], Likang Yao[2], Jiahui Li[2], Nanhao He[2], Siquan Shen[5,6], Yang Li[7], Willem van Schaik[3], Alan McNally[3], Mark A. Holmes[1,9] ✉ & Chao Zhuo[2,9] ✉

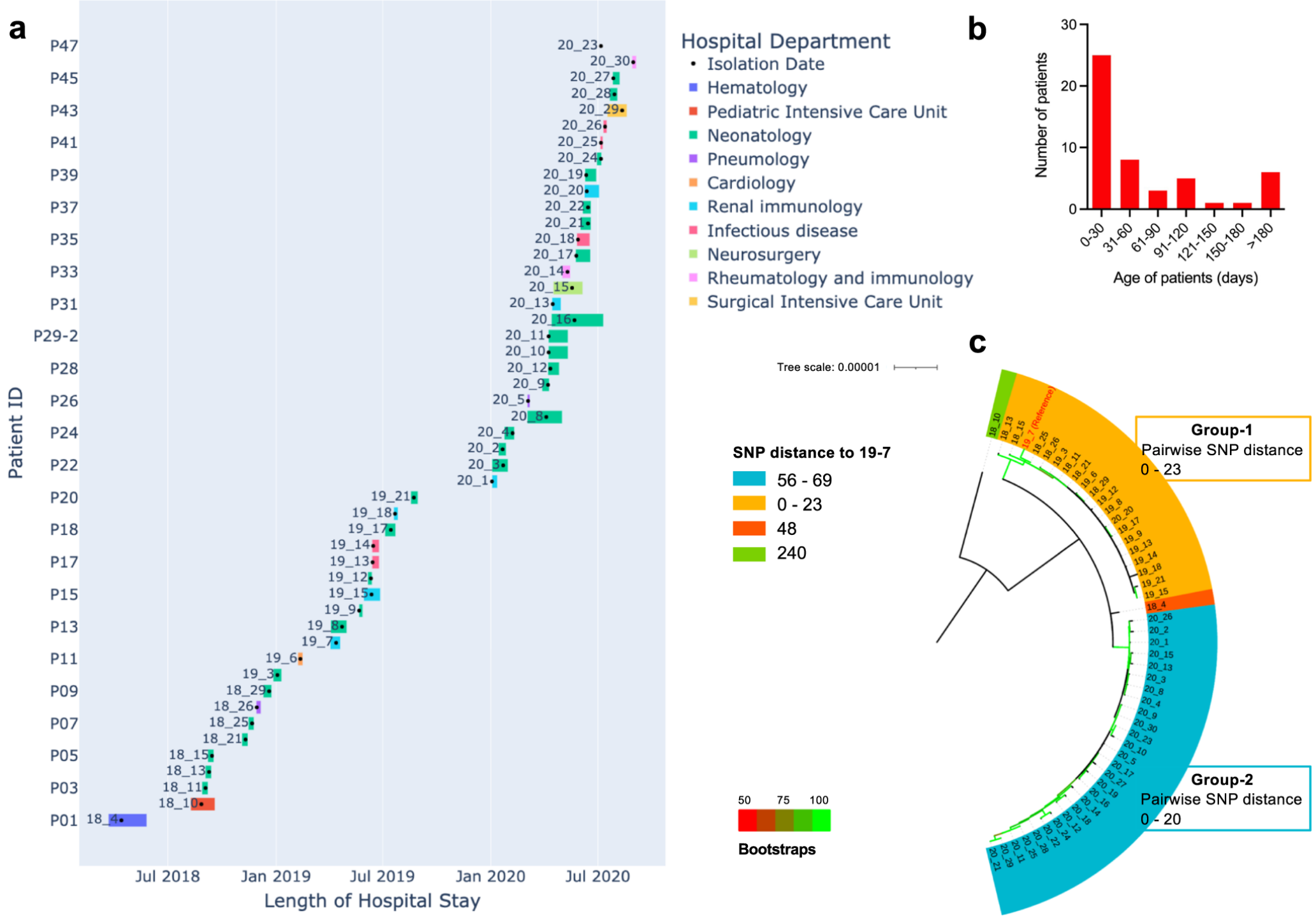
1Department of Veterinary Medicine, University of Cambridge, Cambridge, United Kingdom. 2State Key Laboratory of Respiratory Disease, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, China. 3Institute of Microbiology and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham B15 2TT, United Kingdom. 4Department of Pulmonary and Critical Care Medicine, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China. 5Institute of Antibiotics, Huashan Hospital, Fudan University, Shanghai, China. 6Key Laboratory of Clinical Pharmacology of Antibiotics,MinistryofHealth,Shanghai,China.7DepartmentofClinicalLaboratory,Children'sHospitalofSoochowUniversity, Suzhou,Jiangsu,China.8These authors contributed equally: Xiaoliang Ba, Yingyi Guo. 9These authors jointly supervised this work: Mark A. Holmes, Chao Zhuo. e-mail:mah1@cam.ac.uk; chaosheep@sina.com




## Carbapenem-resistant *Escherichia coli* (CREC) ST410

- CREC prevalence increased in Chinese hospitals between 2017 and 2021.
- ST410 is the most frequent. Four groups of *E. coli* ST410 were identified in this children's hospital,


Number of CREC isolates
120
100
80
60
40
20
0
109
41
12
12
11
11
10
10
10
10
410
167
131
617
156
405
10
101
457
69
Top 10 most isolated STs
Number of CREC isolates
300
250
200
150
100
50
0
4
1
42
3
282
4
2
1
1
1
KPC-2
KPC-19
NDM-1
NDM-4
NDM-5
NDM-9
NDM-13
NDM-20
OXA-48
OXA-181
Carbapenemases
MIC (µg/ml)
>128
128
>64
64
>32
32
16
8
4
2
1
0.5
0.25
≤0.125
≤0.03
0
Imipenem
Meropenem
Ertapenem


a
Patient ID
P47
P45
P43
P41
P39
P37
P35
P33
P31
P29-2
P28
P26
P24
P22
P20
P18
P17
P15
P13
P11
P09
P07
P05
P03
P01
Jul 2018
Jan 2019
Jul 2019
Jan 2020
Jul 2020
Length of Hospital Stay
Hospital Department
Isolation Date
Hematology
Pediatric Intensive Care Unit
Neonatology
Pneumology
Cardiology
Renal immunology
Infectious disease
Neurosurgery
Rheumatology and immunology
Surgical Intensive Care Unit
b
Number of patients
0-30
31-60
61-90
91-120
121-150
150-180
>180
Age of patients (days)
c
Tree scale: 0.00001
SNP distance to 19-7
56 - 69
0 - 23
48
240
50
75
100
Bootstraps
Group-1
Pairwise SNP distance
0 - 23
Group-2
Pairwise SNP distance
0 - 20

## Carbapenem-resistant *Escherichia coli* (CREC) ST410

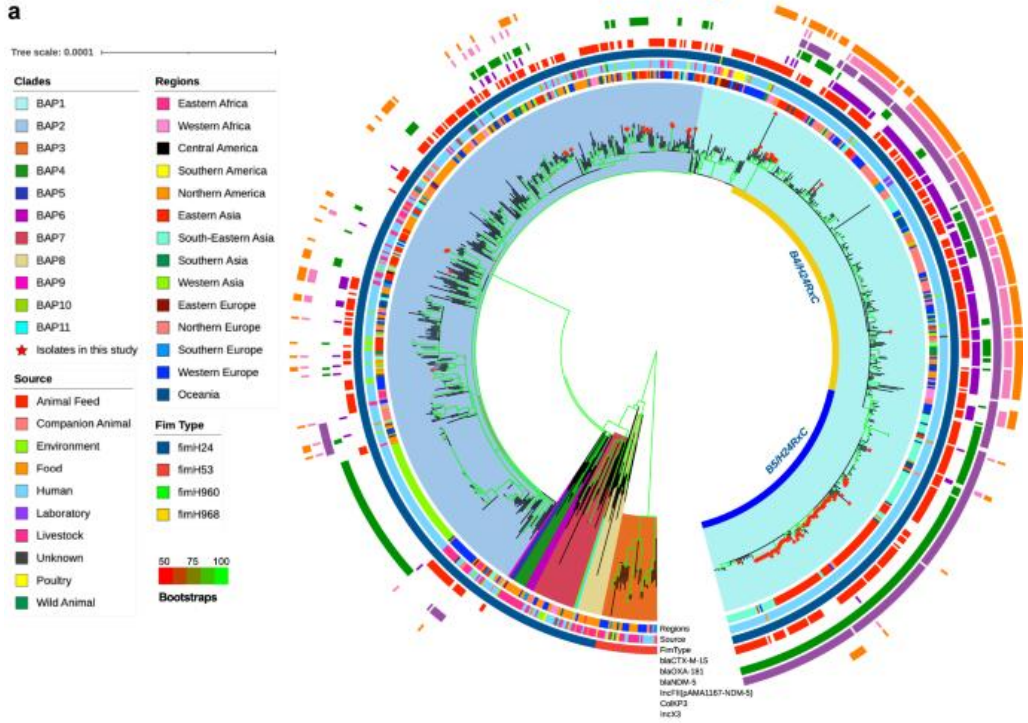
- **Genomic analysis identifies a hypervirulent CREC ST410 clone, B5/H24RxC**

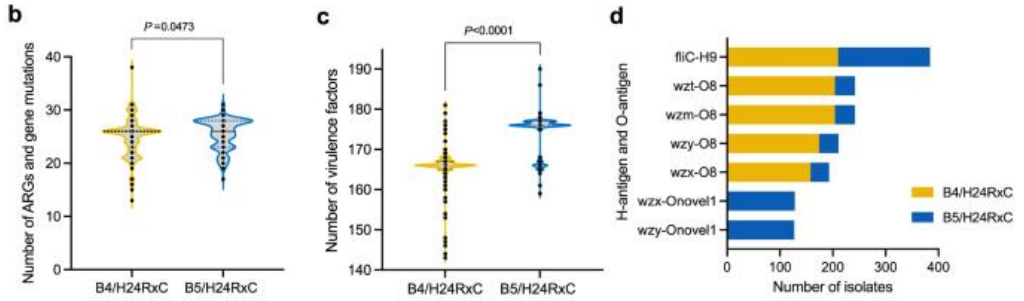
- It may have emerged from the previously characterized B4/H24RxC in 2006

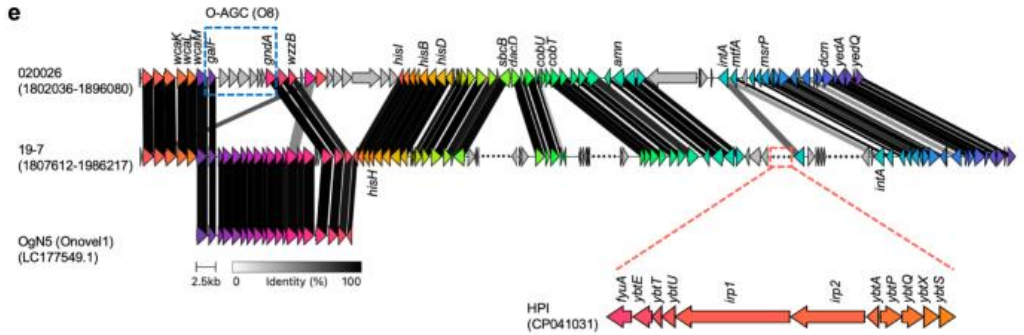
- Compared with B4/H24RxC, **B5/H24RxC lacks the *bla*$_{\text{OXA-181}}$-bearing X3 plasmid, but carries a F-type plasmid containing *bla*$_{\text{NDM-5}}$**
- **Most of B5/H24RxC also carry a high pathogenicity island YBT and a novel O-antigen gene cluster**

- B5/H24RxC grew faster *in vitro* and is more virulent *in vivo*

- Globally disseminated hypervirulent CREC clone, highlights the ongoing evolution of ST410 towards increased resistance and virulence.



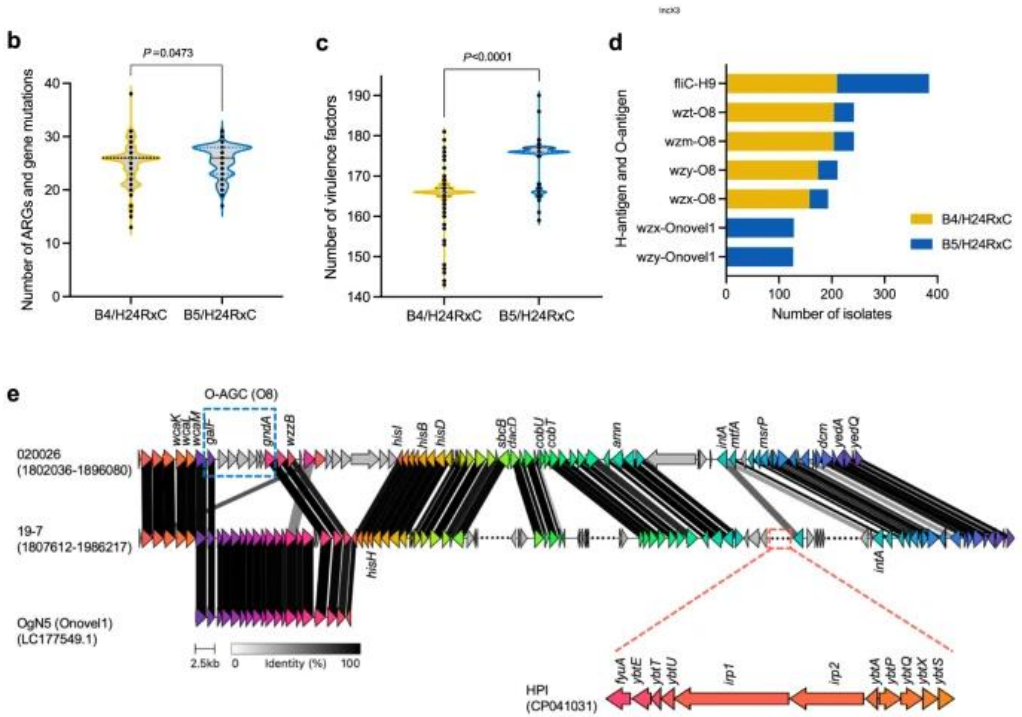

**a** Midpoint rooted maximum-likelihood phylogeny of 956 global ST410 was constructed using a core-genome SNP alignment generated by Snippy v4.6.0 with ST410 isolate YD786 (GenBank accession CP013112.1) as the reference.

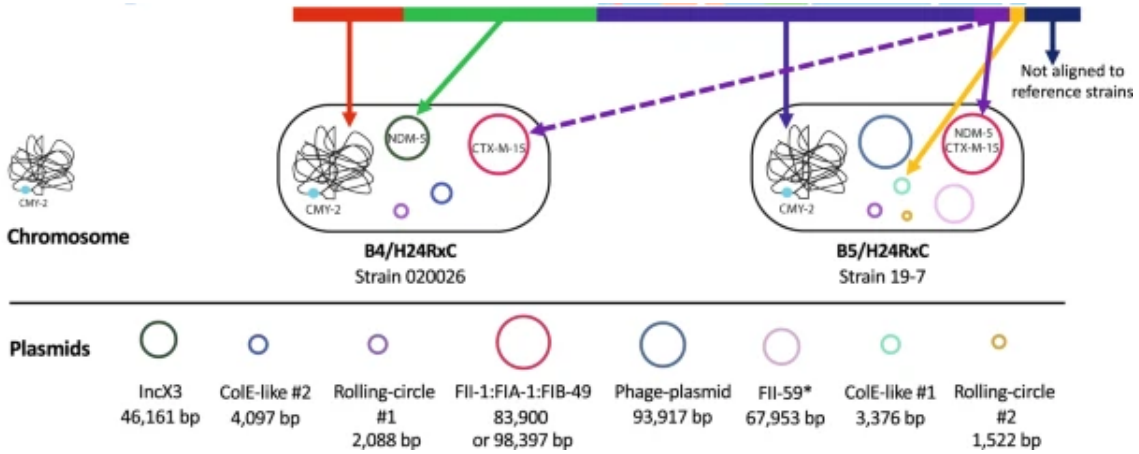
**b** Violin plot showing the distribution of total number of ARGs and mutations that confer resistance in B4/H24RxC and B5/H24RxC clones.

**c** Violin plot showing distribution of total number of virulence factors in B4/H24RxC and B5/H24RxC clones.

**d** Bar plot showing the presence of the lipopolysaccharide (O) and flagellar (H) surface antigens in B4/H24RxC and B5/H24RxC clones.

**e** Comparison of the recombination regions in strain 020026 and 19-7 identified the O-antigen switch from O8 in B4/H24RxC to Onovel1 (OgN5) in B5/H24RxC and the HPI gene cluster in B5/H24RxC clone.



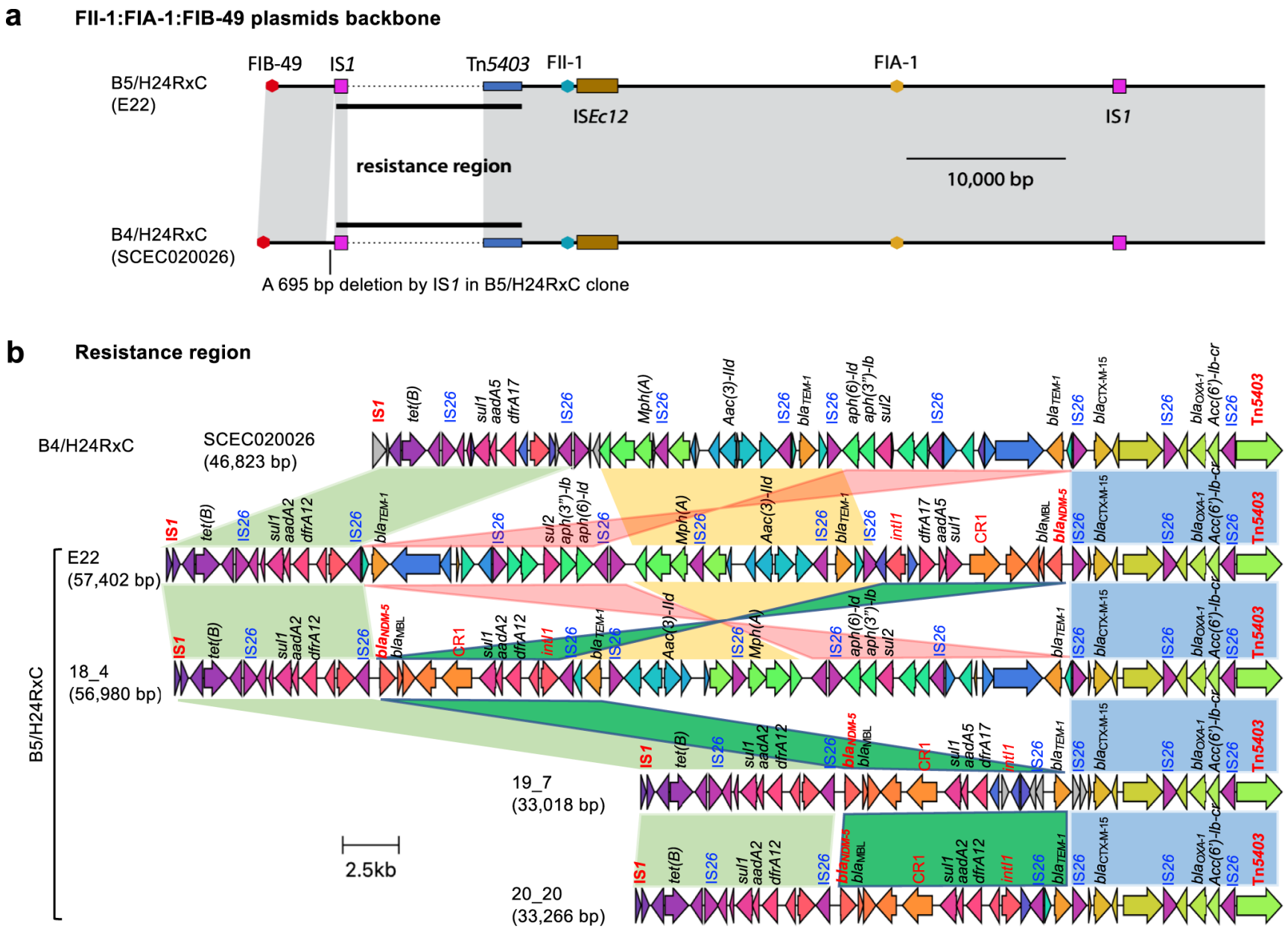

B5/H24RxC had 176 putative virulence genes, B4/H24RxC had 166

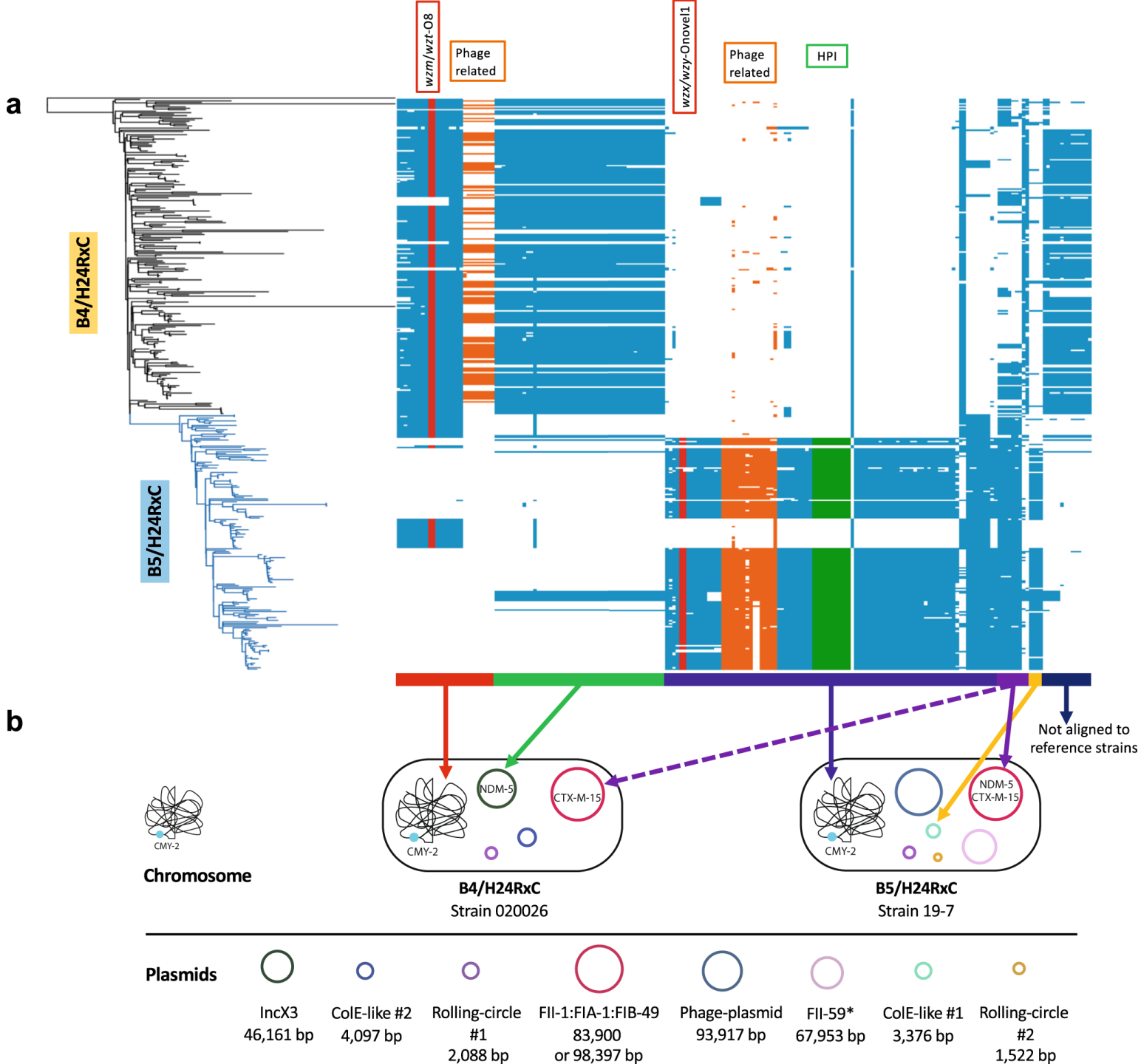
The difference:

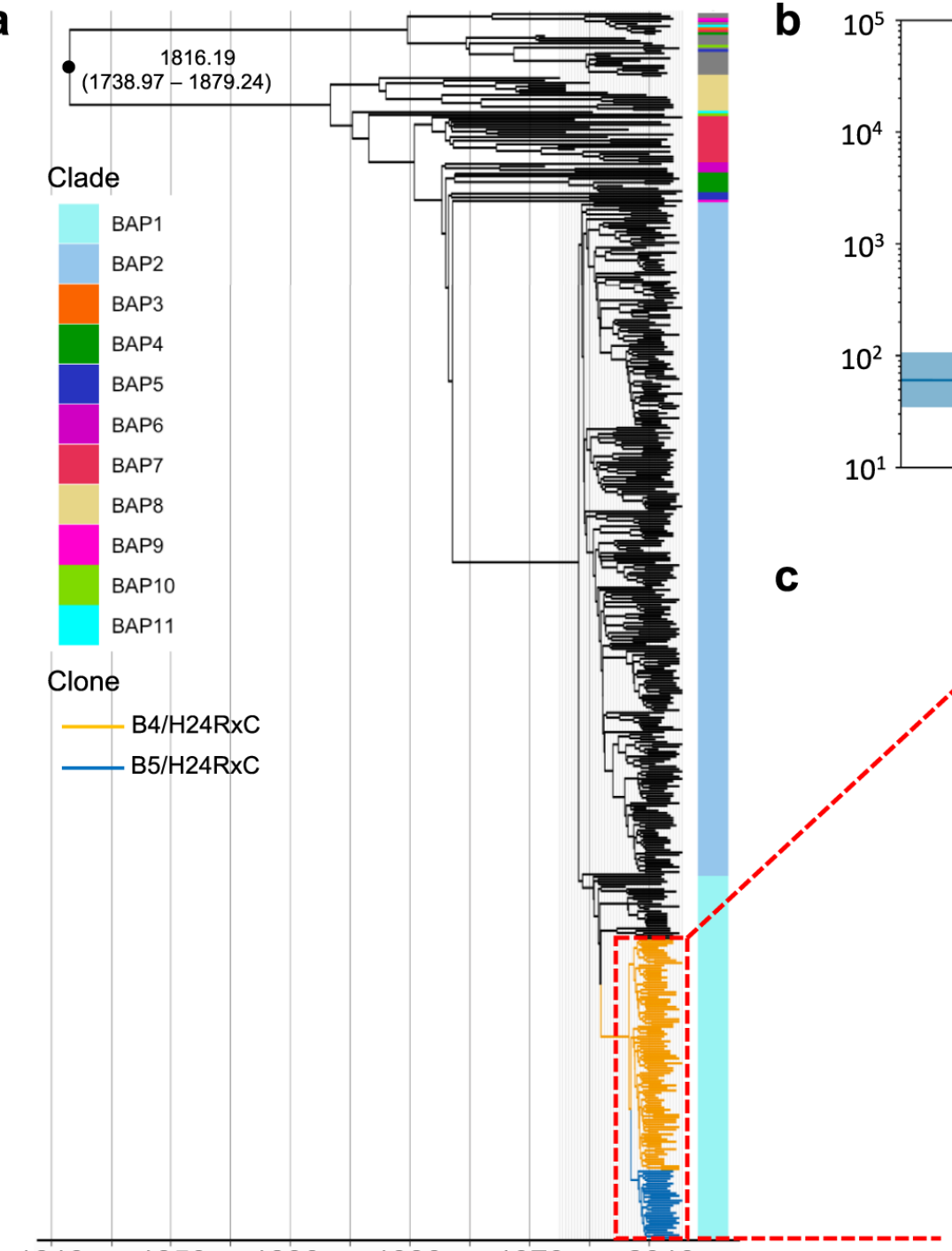
the presence of the **high pathogenicity island (Yersiniabactin)** originated from in Yersinia enterocolitica (*fyuA, irp1, irp2, ybtA, ybtE, ybtP, ybtQ, ybtS, ybtT, ybtU,* and *ybtX)*

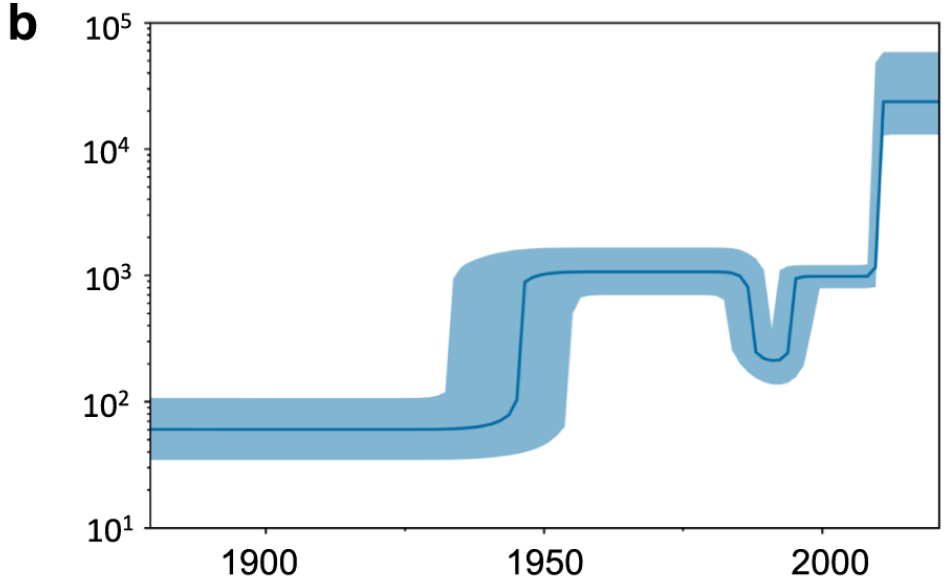
And the O-antigen genes were associated with B4/H24RxC (*wzm*/*wzt* O8) **changed with *wzx/wzy* Onovel1 in B5/H24RxC**

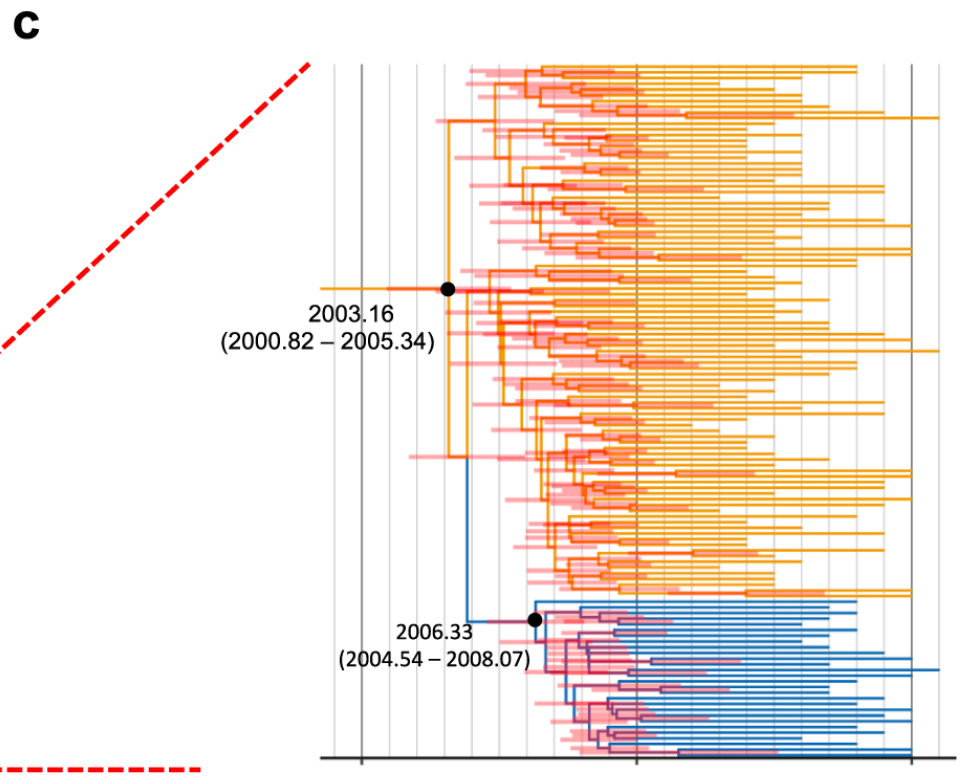
# Emergence of the B5/H24RxC MDR clone driven by recombination and horizontal gene transfer




a
wzm/wzt-O8
Phage related
wzx/wzy-Onovel1
Phage related
HPI
B4/H24RxC
B5/H24RxC
b
Not aligned to reference strains
NDM-5
CTX-M-15
CMY-2
Chromosome
B4/H24RxC
Strain 020026
B5/H24RxC
Strain 19-7
Plasmids
IncX3
46,161 bp
ColE-like #2
4,097 bp
Rolling-circle #1
2,088 bp
FII-1:FIA-1:FIB-49
83,900
or 98,397 bp
Phage-plasmid
93,917 bp
FII-59*
67,953 bp
ColE-like #1
3,376 bp
Rolling-circle #2
1,522 bp




The analysis estimated the age of the ST410 lineage to be approximately 205 years, around 1816 (Fig. 6a)

The B4/H24RxC ancestor was estimated to have originated in 2003 (95% HPD, 2000–2005).

The time of the most recent common ancestor (TMRCA) of B5/H24RxC was estimated at around May 2006 (Fig. 6c).