MICROBIOLOGY

SHORT COMMUNICATION
Le *et al.*, *Microbiology* 2022;168:001232
DOI 10.1099/mic.0.001232

MICROBIOLOGY SOCIETY
OPEN ACCESS

A large chromosomal inversion affects antimicrobial sensitivity of *Escherichia coli* to sodium deoxycholate

Vuong Van Hung Le[1,2]†, Rayén Ignacia León-Quezada[1], Patrick J. Biggs[1,3] and Jasna Rakonjac[1,2,*]

Abstract

Resistance to antimicrobials is normally caused by mutations in the drug targets or genes involved in antimicrobial activation or expulsion. Here we show that an *Escherichia coli* strain, named DOC14, selected for increased resistance to the bile salt sodium deoxycholate, has no mutations in any ORF, but instead has a 2.1 Mb chromosomal inversion. The breakpoints of the inversion are two inverted copies of an IS*5* element. Besides lowering deoxycholate susceptibility, the IS*5*-mediated chromosomal inversion in the DOC14 mutant was found to increase bacterial survival upon exposure to ampicillin and vancomycin, and sensitize the cell to ciprofloxacin and meropenem, but does not affect bacterial growth or cell morphology in a rich medium in the absence of antibacterial molecules. Overall, our findings support the notion that a large chromosomal inversion can benefit bacterial cells under certain conditions, contributing to genetic variability available for selection during evolution. The DOC14 mutant paired with its isogenic parental strain form a useful model as bacterial ancestors in evolution experiments to study how a large chromosomal inversion influences the evolutionary trajectory in response to various environmental stressors.

DATA SUMMARY

All the sequence files in this study are available from the National Center for Biotechnology Information database under the BioProject PRJNA587532.

INTRODUCTION

Genetic variability is essential for bacteria to adapt to changed living environments, such as nutritional availability [1], antimicrobial stresses [2], the presence of bacteriophage predators [3] or competing/cooperating microbial species [4]. Bacteria gain genetic variability by mutations of chromosomal DNA, including nucleotide substitutions, deletions, insertions, gene duplication/amplification and chromosomal inversions or by acquiring new genetic material through horizontal DNA transfer. The power and low cost of next-generation sequencing [5] has enabled large-scale detection of genetic mutations causing altered phenotypes in vast areas of microbiology, including the field of antimicrobial resistance evolution [6–8]. Nonetheless, the data generated from next-generation sequencing, with the predominance of the paired-end short-read Illumina platform, is not sufficient to solve chromosomal inversions caused by homologous recombination of two large, inverted repeats. Examples of such repeats include insertion-sequence (IS) elements, pathogenicity islands, rRNA operons, *rhs* loci, *hsdM-hsdS* operons and prophage [1, 9–13]. Orthogonal techniques are required to identify this class of mutations. Those are either laborious molecular methods such as PFGE or optical mapping of restriction-endonuclease-digested genomic DNA followed by site-specific PCR covering the breakpoints of the inversion for validation [14, 15], or long-read sequencing that incurs a significant extra cost [11]. Therefore, the effect of chromosomal inversions on bacterial physiology has often been overlooked in bacterial evolution studies due to the failure to detect the inversion with short-read sequencing alone. Even when an inversion can be detected, another layer of difficulty is the

Received 27 March 2022; Accepted 08 July 2022; Published 12 August 2022
Author affiliations: [1]School of Natural Sciences, Massey University, Palmerston North, New Zealand; [2]Maurice Wilkins Centre, University of Auckland, Auckland, New Zealand; [3m]EpiLab, Infectious Disease Research Centre, School of Veterinary Science, Massey University, Palmerston North, New Zealand.
***Correspondence:** Jasna Rakonjac, j.rakonjac@massey.ac.nz
Keywords: antimicrobial susceptibility; chromosomal inversion; hybrid genome assembly; IS5 element; sodium deoxycholate.
Abbreviations: AMP, ampicillin; CIP, ciprofloxacin; DOC, sodium deoxycholate; IS, insertion-sequence; MEM, meropenem; ORF, open reading frame; PFGE, pulsed-field gel electrophoresis; VAN, vancomycin.
†Present address: Section of Microbiology, Department of Biology, University of Copenhagen, Denmark.
Five supplementary figures and two supplementary tables are available with the online version of this article.

co-occurrence with numerous small mutations, typically seen in long-term evolution experiments or amongst natural isolates. The combination of chromosomal inversion with superimposed point mutations makes it unfeasible to attribute the altered phenotype to the former or the latter type of genetic changes [1, 16–18].

In a previous study, we investigated the possibility and mechanism for a TolC-associated-efflux-pump-deficient *Escherichia coli* strain to gain resistance to sodium deoxycholate (DOC), a secondary bile salt in the human gut, that is important for fat digestion and the balance of the gut microbiota [19]. Using short-read genomic sequencing of the selected mutants having decreased sensitivity to DOC, we were able to identify mutations in all, except one mutant, named DOC14, that had no detectable changes in DNA sequence [19]. In this report, we describe molecular, bioinformatic and genomic approaches that revealed a large chromosomal inversion of about 2.1 Mb in length, mediated by homologous recombination between two IS*5* sequences. We show that this inversion is responsible for changes in interactions with multiple antibacterials: decreased sensitivity to DOC, ampicillin and vancomycin, and increased susceptibility to ciprofloxacin and meropenem.

METHODS

Bacterial strains, growth conditions and antimicrobial agents

The *E. coli* strain JW5503 (BW25113 Δ*tolC*) was derived from the *E. coli* K-12 single-gene knock-out Keio collection [20], while the DOC14 mutant, designated K2635 [JW5503 IN(*insH2-insH9*)], was selected for decreased DOC susceptibility of the Δ*tolC* parental strain in our previous work [19]. *E. coli* cultures were grown in 2×YT medium (BD Difco) at 37 °C with shaking at 200 rpm. DOC was provided by New Zealand Pharmaceuticals. Ampicillin, ciprofloxacin and vancomycin were purchased from Gold Biotechnology. Meropenem was purchased from AK Scientific.

Bioinformatic analysis of illumina short-read sequences

Bacterial genomic DNA was extracted from cultures in the exponential growth phase (OD_{600nm} of ~0.2) using the UltraClean Microbial DNA Isolation Kit (Qiagen). These early exponential-phase cultures were obtained by diluting an overnight culture (~5×10^9 c.f.u. ml^{-1}) 100-fold with fresh 2×YT medium and incubating the diluted cultures with 200 r.p.m. shaking for 1.5 h at 37 °C until they reached an OD_{600nm} of ~0.2. Purified genomic DNA samples were then submitted to the Massey Genome Service (Massey University, Palmerston North, New Zealand) for whole genome sequencing using an Illumina TruSeq Nano DNA library preparation and 2×250 base paired-end Nano v2 sequencing on the Illumina MiSeq platform. The raw reads were trimmed to a quality cut-off value of Q30 and the short-length reads (<25 bases by default) were removed using SolexaQA^{++} v3.1.7.1 [21]. The trimmed reads were aligned with the *E. coli* BW25113 reference genome (GenBank accession number CP009273.1) or the newly assembled parental genome JW5503 using Bowtie2 v2.4.2 with --very-sensitive mode [22]. The resulting *.sam* alignment file was then converted to *.bam* format and sorted, from which the genome coverage depth at individual coordinates was extracted using Samtools v1.14 [23]. The depth of genome coverage was visualized in the R environment v4.0.5 [24]. Genetic variants were called from the *.bam* file using freebayes v1.3.1 (-p 1 --min-alternate-fraction 0.5) [25] and functionally annotated using SnpEff v5.0e [26]. The same analysis was applied to the Illumina short-read data obtained from the stationary-phase overnight cultures in our previous study [19].

To analyse the biphasic nature of the 1839-base invertible element within the e14 prophage, the DNA fragment containing 1839-base invertible sequence and two 1000-nt flanking sequences was extracted from the assembled JW5503 genome sequence (coordinates: 1202788–1206626) as the forward orientation (Fig. S2A) (available in the online version of this article) and the corresponding DNA fragment having the 1839-base invertible sequence reverse-complemented was prepared in the reverse orientation. Mapping of Illumina short reads to these two template sequences and visualization of the read coverage depth was done as described above.

Nanopore long-read sequencing and consensus hybrid genome assembly

The genomic DNA of the DOC14 mutant and its parental strain JW5503 was extracted from overnight cultures using the UltraClean Microbial DNA Isolation Kit (Qiagen). The sequencing library was prepared from 1 µg of genomic DNA using the Oxford Nanopore Ligation Sequencing Kit (SQK-LSK109) according to the manufacturer's instructions and sequencing was done for 12 h for each genome library on a FLO-MIN106 R9.4.1 flow cell with the MinION Mk1B device. Each run generated about 400 000–456,000 reads, estimated to cover the genome of the *E. coli* strains in this study (~4.6 Mb) to an average depth of more than 950. The raw reads were base-called using DeepNano-blitz at high-accuracy mode (--network-type 96) [27] and the adapters were then trimmed using Porechop v0.2.4 (https://github.com/rrwick/Porechop). Reads that had an average read quality score <10 or were shorter than 1000 bases were removed using NanoFilt v2.7.1 [28]. The genome was assembled using the consensus long-read-first hybrid approach as recommended by the Trycycler developer [29]. Briefly, the long-read data set for each strain was randomly split into 12 read groups that were used to assemble 12 independent genomes using three different assemblers, including Flye v2.9 [30], miniasm v0.3-r179+miniPolish v.v0.1.3 [31, 32] and Raven v1.6.1 [33] using default parameters. Each assembler contributed three assembled genomes. The consensus assembly was then obtained from the 12 independently assembled

genomes using Trycycler v0.5.1 [29], which was then polished with long reads using Medaka v1.4.4 (https://github.com/nanoporetech/medaka) and with pooled short reads (combining reads from exponential phase and overnight cultures) using Pilon v1.24, PolyPolish v0.4.3 [34] and Polca from the toolkit MaSuRCA v4.0.5 with default parameters [35]. Whole genome alignment was performed using the progessiveMauve algorithm [36] and visualized using Mauve v2.4.0 [37]. Mapping of long reads to the newly assembled parental genome JW5503 was done using minimap2 v2.17 with default parameters [38] and visualized using the Integrative Genomic Viewer v2.11.9 [39]. All the sequence data in this study were submitted to the National Center for Biotechnology Information database under the BioProject PRJNA587532 (Table S1).

Three-primer PCR assays

Genomic DNA for PCR analysis was extracted from bacterial colonies using the rapid boiling method as previously described [40]. To probe the flanking region of the IS*5* loci in the DOC14 mutant and the parental strain, PCR assays were performed with the DreamTaq Hot Start DNA Polymerase (Thermo Scientific) according to the manufacturer's instructions. The primers used are listed in Table S2, which were designed using Primer-Blast [41]. The PCR products were analysed by agarose gel electrophoresis and submitted for Sanger sequencing using the BigDye Terminator v3.1 chemistry at the Massey Genome Service (Massey University).

Antimicrobial sensitivity assays

The MIC for each antimicrobial agent was determined using a standard agar dilution assay as previously described [42]. For antimicrobial susceptibility spot plating assays, three independent overnight cultures of each strain were prepared at 37 °C, followed by 10-fold serial dilutions (10^1 to 10^8) of these cultures with 2×YT medium. Next, 5 µl of each dilution was spotted on 2×YT agar plates containing DOC (100 or 75 µg ml^{-1}), ampicillin (6 µg ml^{-1}), ciprofloxacin (0.00195 µg ml^{-1}), vancomycin (400 µg ml^{-1}) or meropenem (0.03125 µg ml^{-1}) or without antimicrobial agents. The agar plates were incubated at 37 °C for 24 h before examining bacterial growth.

Transmission electron microscopy

Bacterial cultures were grown overnight with 2×YT at 37 °C. A carbon-coated 200 mesh copper grid (Agar Scientific) was floated in a volume of 50 µl of the cell culture for 3 min. The grid was then rinsed with MilliQ water for 1 min followed by being negatively stained with a 20 µl drop containing 3% ammonium molybdate for 4 min. The grid was then rinsed again with MilliQ water for 1 min and allowed to dry before observation. All the micrographs were taken using a 100 kV FEI Tecnai G2 Spirit BioTWIN transmission electron microscope at the Manawatu Microscopy and Imaging Centre (Massey University).

Bacterial growth measurement

Bacterial growth was performed in 2×YT medium in 96-well microtitre plates. Each exponential culture comprising 200 µl of 10^6 c.f.u. ml^{-1} was incubated at 37 °C without shaking and the optical density at 600 nm was monitored every 20 min for 24 h. Twenty-four replicates were included for the parental strain and the DOC14 mutant. The growth curve was visualized using the R package ggplot2 v3.3.3 in the R environment v4.0.5 [24, 43].

RESULTS

Short-read sequencing to identify DNA mutations in the *E. coli* DOC14 mutant

In our previous study, we isolated 20 mutants that were selected for a decreased susceptibility to DOC of a Δ*tolC* efflux-pump-deficient *E. coli* parental strain. All mutants had an identical MIC, determined by the agar dilution assay, of 125 µg ml^{-1}, slightly higher than that of the parental strain (100 µg ml^{-1}) [19]. By mapping the Illumina short-reads derived from the overnight cultures to the BW25113 reference genome, we found DNA mutations in all except one mutant, called DOC14 (Table 1) [19]. This suggests that the underlying mutation in DOC14 may be a large-scale chromosomal structural rearrangement such as a duplication and/or an inversion. The former was ruled out by examining the read depth coverage along the genome in the stationary phase cultures, showing no unusual jumps in the coverage depth (Fig. 1a, b).

We then sequenced the genomic DNA extracted from the exponentially growing cultures of the parental strain and the DOC14 mutant and compared the read coverage depth along the genome when mapping the reads to the BW25113 reference genome. The rationale is that under the exponential phase, the DNA copy number is greater in the region near the origin of replication (*oriC*) than in the region closer to the terminus due to a replication-dependent gene dosage effect, rendering an inverted-V shape of read coverage depth distribution, with the central peak at the *oriC* [44, 45]. If any inversion occurs in the mutant chromosome, this coverage depth distribution will be flipped in the inverted portion of the genome. Surprisingly, the read depth coverage of both strains appeared to have an irregular pattern, deviating from the expected inverted V-shape coverage distribution (Fig. 1c). Nonetheless, we found that in the DOC14 mutant, a region spanning the coordinates of 0.6–1.6 Mb had higher coverage than expected and a region spanning the coordinates of 1.8–3.1 Mb had lower coverage than expected, relative to the genome coverage

Table 1. Differences in the genomic sequences of the parental strain JW5503 and the DOC14 mutant relative to the reference genome BW25113 (GenBank accession no. CP009273.1)*

Genetic differences†	JW5503	DOC14
Deletion	In-frame deletion of the *tolC* gene‡	In-frame deletion of the *tolC* gene‡
Missense mutation	C4531989T [Leu46Phe in the FimE protein]	C4531989T [Leu46Phe in the FimE protein]

*Mapping the Illumina short reads from overnight cultures or exponentially growing cultures gave the same results, as listed in the table.
†The parental strain JW5503 and the DOC14 mutant had the same mutations relative to the reference genome BW25113, indicating that there is no genetic difference (i.e. substitutions, deletions, insertions) between the genomes of these two strains that can be identified by mapping Illumina short reads to the reference genome.
‡The genetic make-up of the parental strain JW5503 (BW25113 Δ*tolC*).

depth pattern of the parental strain (Fig. 1d). This observation suggests that there is an inversion spanning the coordinates of 0.6–3.1 Mb in the DOC14 mutant, although we could not identify the exact breakpoints of the chromosomal inversion with the Illumina short reads alone.

Hybrid genome assembly to identify the chromosomal inversion in the DOC14 mutant

We sequenced the genomic DNA of the DOC14 mutant and its parental strain using the Oxford Nanopore sequencing platform and assembled their genomes using a consensus long-read-first hybrid approach (see Methods, Fig. 2a) [29]. This gave rise to a single circular chromosome of 4630639 nt in length for both *E. coli* strains. Performing pairwise genome alignment by progressiveMauve [36] indicates a C221333T substitution within the 16S rRNA-encoding gene *rrsH*, a 1839-base inversion within the e14 prophage and a 2.1 Mb chromosomal inversion limited by IS*5* P2 and P9 sequences in the DOC14 mutant relative to the parental strain (Fig. 2b). However, by mapping the Illumina short reads to the assembled parental genome JW5503 and manually examining the read alignment, we confirm that the nucleotide at the 221333rd coordinate was **T** in the parental strain and the

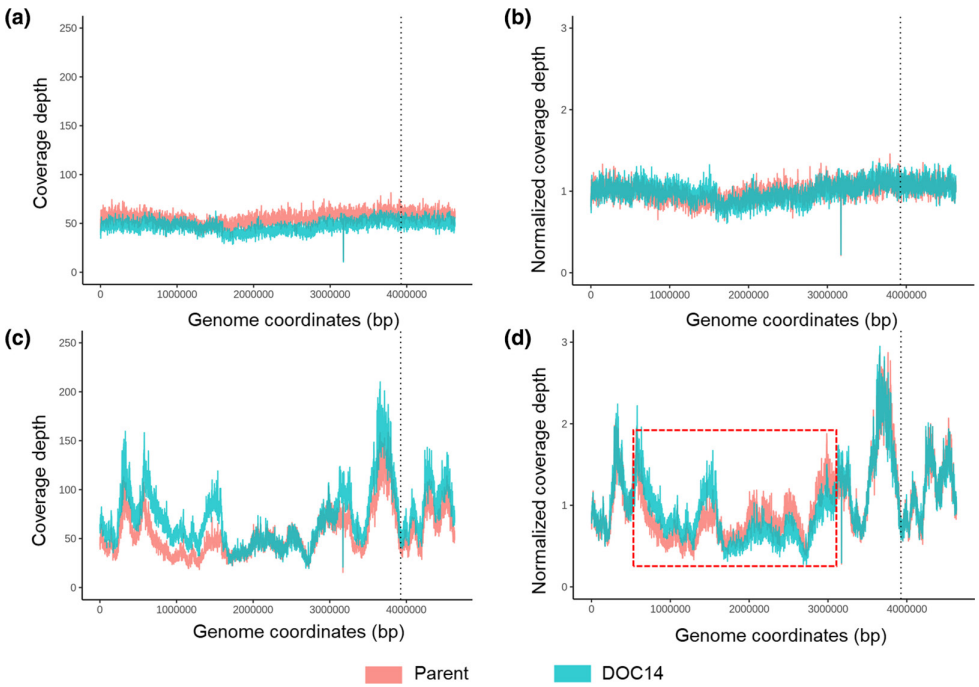

Fig. 1. Genome read coverage depth of the parental strain JW5503 and the DOC14 mutant. (a) Absolute coverage depth, stationary phase; (b) coverage depth normalized by mean depth, stationary phase; (c) absolute coverage depth, exponential phase; (d) coverage depth normalized by mean depth, exponential phase. Genomic DNA was extracted from the overnight culture (a and b) or exponentially growing cultures (c and d) of the parental strain (salmon) and the DOC14 mutant (cyan). The Illumina short reads were aligned to the *E. coli* BW25113 reference genome (GenBank accession no. CP009273.1). The read coverage depth was plotted using a window size of 2000 nt. The black vertical dotted line indicates the chromosomal origin of replication *oriC*. The red dashed box highlights the region having a difference in the genome read coverage depth between the parental strain JW5503 and the DOC14 mutant.

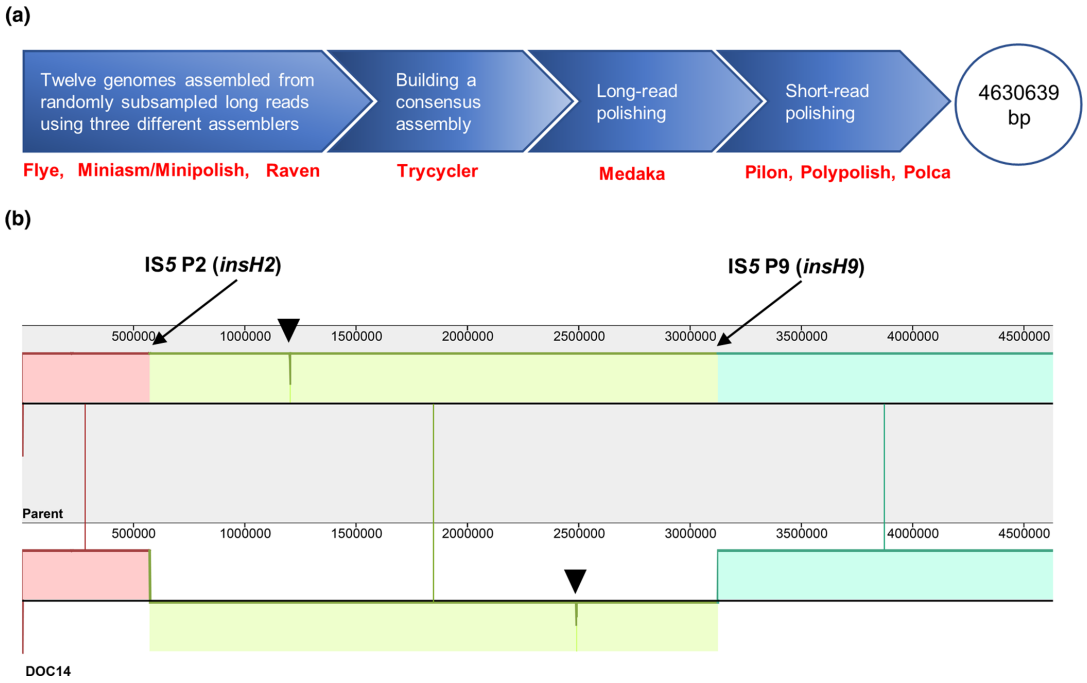

Fig. 2. Hybrid genome assembly to identify the chromosomal inversion in DOC14. (a) Block-arrow scheme; the workflow for *de novo* genome assembly of the parental strain and the DOC14 mutant using a consensus long-read-first hybrid approach. Text in red below the block-arrow scheme refers to the software used at each stage. (b) Genome alignment of the parental strain JW5503 and the DOC14 mutant. Locally collinear blocks are colour-coded. Sequence orientation relative to the reference genome: above the centre line, forward orientation; below the centre line, reverse orientation. The strand shift in the DOC14 genome represents a chromosomal inversion mediated by homologous recombination of the IS*5* P2 and P9 sequences (containing the transposase-encoding genes *insH2* and *insH9*, respectively). The black triangle indicates a 1839-nt inversion within the prophage e14 (coordinates: 1203788–1205626). Sequence alignment was processed and visualized using Mauve software, setting the parental strain JW5503 as the reference [37].

DOC14 mutant (Fig. S1) and the C221333T substitution was not a *de facto* mutation. Instead, such a genetic difference that is in a gene existing in multiple copies (seven 16S rRNA genes per *E. coli* K-12 genome) is probably an artefact arising from incomplete genome polishing with Illumina short reads in the genome assembly/polishing pipeline. A 1839-base inversion that is also detected corresponds to an invertible region within the e14 prophage, which has been reported to exist as a biphasic variation in an *E. coli* K-12 population (Fig. S2a) [46]. Here, by mapping the Illumina short reads to two possible forward and reverse configurations of this region, we show that both the parental strain and the DOC14 mutant had a biphasic configuration in this region (Fig. S2b). Thus, the observed 1839-base inversion is also not a true genetic difference, but simply results from a random orientation of each configuration by the genome assembler. This leaves the 2.1 Mb inversion limited by the two IS*5* elements (IS5 P2: 570413–571607 and P9: 24052–3125246) as the only mutation in the DOC14 mutant, which would subsequently explain any phenotypic differences between the DOC14 mutant and its parental strain.

We further confirmed the chromosomal inversion in the DOC14 mutant using three-primer PCR assays to probe the region flanking the breakpoints of the inversion, followed by sequencing the PCR products (Fig. S3). In addition, when mapping the Nanopore long reads to the assembled parental genome JW5503 using minimap2 [38], we found an absence of DOC14-derived reads that can map across the two genes flanking the inversion breakpoints at a coverage depth of ~700 (Fig. S4), indicating that a reversion of the 2.1 Mb inversion to restore the original conformation, if any, would occur at a very low frequency in the DOC14 population (<1:700). The inversion breakpoint loci as annotated in the *E. coli* MG1655 genome sequence are the transposase-encoding genes *insH2* (in the IS5 P2) and *insH9* (in the IS5 P9); this inversion is therefore designated IN(*insH2- insH9*).

The IN(*insH2-insH9*) inversion in the DOC14 mutant affects antimicrobial sensitivity without causing a significant defect in bacterial growth or cell morphology

As previously mentioned, the DOC14 mutant was slightly less sensitive to DOC than the parental strain (MIC 125 vs. 100 μg ml^{-1}) in a standard agar dilution assay [19, 42]. To increase the resolution of the antimicrobial susceptibility difference, we used a spot plating assay, in which 10-fold serial dilutions of overnight cultures were plated on the agar plates containing the antimicrobial agents at the concentration near the MIC of the parental strain [47]. As shown in Fig. 3a, the DOC14 mutant had a higher plating efficiency than the parental strain at 0.75 × MIC (75 μg ml^{-1}) and 1 × MIC (100 μg ml^{-1}) of DOC, by 100- and 1000-fold, respectively. Plating efficiency of the parent vs. DOC14 was also tested on ampicillin (6 μg ml^{-1}, 0.75× MIC), meropenem (0.03125

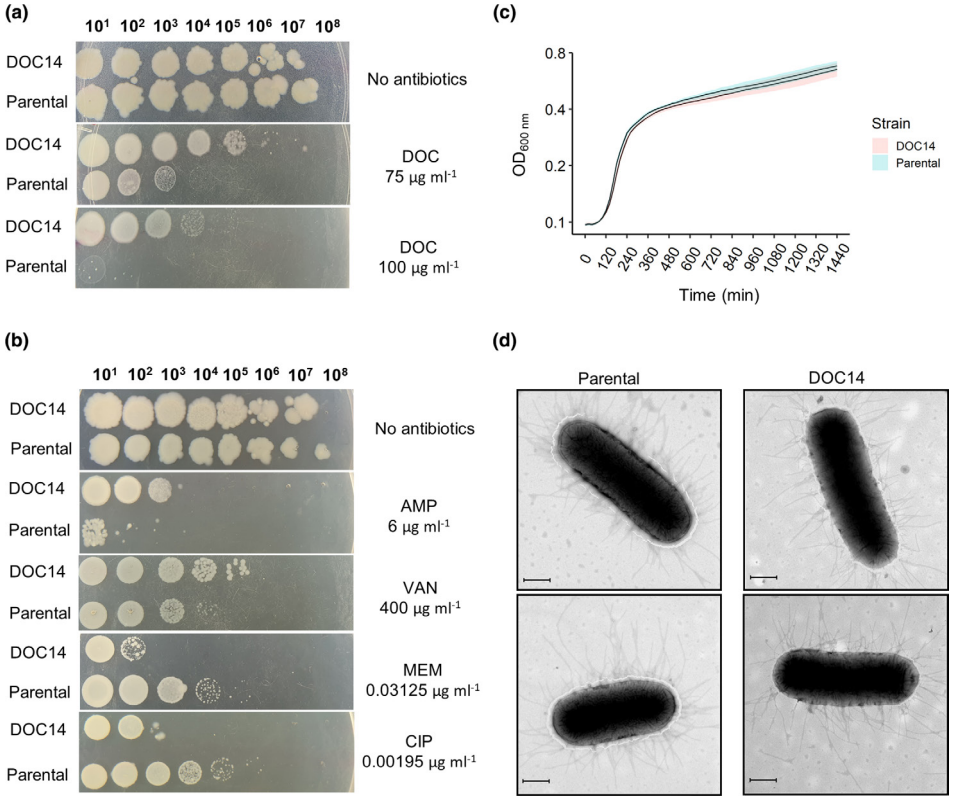

Fig. 3. Effect of the IS*5*-mediated inversion on the phenotype of the DOC14 mutant. Sensitivity of the parental strain (Δ*tolC*) and the DOC14 mutant to sodium deoxycholate (a), and ampicillin, vancomycin, meropenem and ciprofloxacin (b). Representative data from three biological replicates are shown. From left to right, 10-fold serial dilutions of the overnight culture. DOC, sodium deoxycholate; AMP, ampicillin; VAN, vancomycin; MEM, meropenem; CIP, ciprofloxacin. (c). Semi-logarithmic plot of the growth curve of the parental strain and the DOC14 mutant. The optical density at 600 nm ($OD_{600\ nm}$) was measured every 20 min for 24 h. The centre line represents the mean values of 24 culture replicates; the standard deviation is presented by the shaded areas. (d) Representative electron micrograph of the parental strain (left) and the DOC14 mutant (right). Bar, 500 nm.

μg ml^{-1}, 1× MIC), ciprofloxacin (0.00195 μg ml^{-1}, 0.5× MIC) and vancomycin (400 μg ml^{-1}, 0.5× MIC) as representatives of the major antibiotic classes: β-lactams (penicillins and carbapenems), fluoroquinolones and glycopeptides, respectively. The sensitivity of the DOC14 mutant was lower to ampicillin and vancomycin, but higher to meropenem and ciprofloxacin, in comparison with that of the parental strain (Fig. 3b).

To examine whether the chromosomal inversion affects the growth of the DOC14 mutant in 2×YT rich medium, which had been used during mutant selection and antimicrobial susceptibility assays, we monitored the bacterial growth over a 24 h period in the absence of any antimicrobial agents and showed that there was no significant defect in the bacterial growth of the DOC14 mutant in comparison with its parent (Fig. 3c). In addition, the transmission electron micrographs did not show any noticeable difference in cell morphology between DOC14 and its parental strain (Fig. 3d). These observations also rule out the involvement of bacterial growth as a determinant of the change in antimicrobial susceptibility.

DISCUSSION

In addition to genetic variants involving small changes, large chromosomal inversions have been observed as landmark events in bacterial evolutionary history. Classical examples are the inversion that distinguishes between *Salmonella enterica* serovar Typhimurium and *E. coli* K-12, between *S. enterica* serovars Typhimurium and Typhi [16, 17], or between two closely related *E. coli* K-12 wild-type laboratory strains, MG1655 and W3110 [18, 44]. Furthermore, a long-term evolution experiment with *E. coli* propagated in a glucose-limited environment for over 25 years demonstrated the emergence of large chromosomal inversions (from ~164 kb to ~1.8 Mb in size) in independently evolved clones [1]. In other studies, large chromosomal inversions have been detected in multiple clinical isolates of *Pseudomonas aeruginosa* sampled more than 20 years apart in a cystic fibrosis patient [13] and in different lineages of *E. coli* O157:H7, a life-threating human zoonotic pathogen [10]. In those examples, there are, nonetheless, many other genetic differences, leaving an unresolved question regarding whether the identified chromosomal inversions confer any fitness advantage to the bacterial host or whether their co-occurrence with selected beneficial mutations is fortuitous and they are phenotypically neutral. Chromosomal inversions without other confounding mutations in which a causal

link to a phenotype has been determined are very rarely found. Such published examples include a chromosomal inversion in *Staphylococcus aureus* leading to two phenotypic changes, formation of a small-colony variant and increased survival in the whole blood bactericidal assay [11, 12]. Here, we report a large chromosomal inversion (~2.1 Mb) caused by homologous recombination between two oppositely oriented IS*5* sequences that is the sole genetic difference between the DOC14 mutant, designated as K2635 [BW25113 Δ*tolC,* IN(*insH2- insH9*)], and its parental strain JW5503 (BW25113 Δ*tolC*). The identified inversion was shown to confer an improved survival of *E. coli* under exposure to the secondary bile salt DOC, ampicillin and vancomycin, without introducing any significant defect to bacterial growth or cell morphology (Fig. 3). Our findings and other mentioned examples [11, 12] indicate that large chromosomal inversions, although highly constrained, may confer advantages to bacterial cells under certain growth conditions, highlighting a potential adaptive role of this class of mutations in bacterial evolution. The fact that the chromosomal inversion in the DOC14 mutant was obtained in a selection experiment for decreased susceptibility to DOC further supports this suggestion [19].

Of great interest is how such a large inversion, which causes no change in the gene content, can exert an effect on the phenotype, let alone a differential effect on the susceptibility to those antibiotics that target bacterial cell wall synthesis: vancomycin, ampicillin and meropenem (Fig. 3). We propose three hypotheses, based on the genetic context of the IN(*insH2- insH9*) inversion in the DOC14 mutant, of this differential effect on antibiotic sensitivity. First, due to the asymmetric nature of the inversion across *oriC*, the inversion moves the genes located downstream of the IS*5* P2 element in the parent closer to *oriC* in DOC14, while moving the genes upstream of the IS*5* P9 element further away from *oriC*. The change in proximity to *oriC* might affect their expression in the exponential phase of growth, through a replication-dependent gene dosage effect. This is reflected by the difference in coverage depth when mapping the sequencing reads to the reference genome (Fig. 1d, red box). Second, the IS*5* P2 element is located within the pseudo-gene *nmpC*, which is flanked by the *quuD* gene and the cell lysis cassette (*essD*, *ybcS*, *rzpD*/*rzoD*) of the DLP12 defective prophage (Fig. S5). The protein Q^{DLP12}, encoded by the *quuD* gene, positively regulates transcription of the DLP12 lysis cassette located on the opposite flank of the IS*5* P2 [48]. Given this gene arrangement, the chromosomal inversion in the DOC14 mutant would spatially distance the lysis cassette from its regulator-encoding gene *quuD,* thereby modulating the efficiency of regulation. The third consequence of the inversion is splitting of the prophage DLP12 genome into two segments, 2.1 Mb away from each other. Therefore, the inversion in DOC14 prevents DLP12 prophage excision, an event that has been shown to be proficient in the parental chromosomal conformation [49]. The exact manner by which these changes ultimately affect susceptibility to different antibiotics remains to be elucidated (Fig. 3a, b). Comparative transcriptomic and proteomic profiling of the DOC14 mutant and its parent would be the first step in resolving the above hypotheses.

Regarding the read coverage depth pattern in the exponential growth phase, it is intriguing to observe multiple peaks along the chromosome (Fig. 1d). This replication pattern is reminiscent of the *oriC*-independent multi-site DNA replication initiation [50–52]. Since this occurs in both strains (DOC14 and its parent), the multi-peak pattern is not caused by the IN(*insH2- insH9*) inversion. While this is beyond the scope of our study, future work to identify the cause of this phenomenon is warranted.

Overall, we report a combinatorial use of the short-read and long-read sequencing to pinpoint a 2.1 Mb chromosomal inversion of a bile-salt-selected laboratory *E. coli* strain. This inversion, through an as yet unclear mechanism, causes an increase or decrease in the susceptibility to different antibiotics without changing bacterial growth or cell morphology in the antibiotic-free medium. The two isogenic strains (parental and DOC14), differing solely by a chromosomal inversion and containing no additional mutations, are promising model organisms that will allow elucidation of the mechanisms by which a large chromosomal inversion may influence bacterial evolutionary trajectory in response to various growth conditions.

Funding information
This work was supported by a Massey University-MBIE PSAF II grant MU001985, Dextra – New Zealand Pharmaceuticals and a generous donation by Anne and Bryce Carmine. V.V.H.L. was supported by a Callaghan PhD Scholarship NZPX1501 and Massey University School of Natural Sciences. R.I.Q.L. was supported by a PhD scholarship from Anne and Bryce Carmine and Massey University School of Natural Sciences. J.R. was supported by Massey University, School of Natural Sciences. P.J.B. was supported by Massey University School of Natural Sciences and School of Veterinary Sciences. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
We thank Xiao Xiao Lin (Massey Genome Service, Massey University, Palmerston North, New Zealand) for excellent genome sequencing services and Raoul Solomon (Manawatu Microscopy and Imaging Centre, Massey University, Palmerston North, New Zealand) for technical assistance with transmission electron microscopy.

Author contributions
Conceptualization, V.V.H.L. and J.R.; funding and lead, J.R.; realization, V.V.H.L.; bioinformatics, V.V.H.L. and P.J.B.; microscopy, R.I.Q.L.; original draft preparation, V.V.H.L.; review and editing, P.J.B. and J.R.

Conflicts of interest
The author(s) declare that there are no conflicts of interest

Ethical statement
No experiments were performed on humans or animals

References

1. Raeside C, Gaffé J, Deatherage DE, Tenaillon O, Briska AM, *et al*. Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *mBio* 2014;5:e01377-14.
2. Hoeksema M, Brul S, Ter Kuile BH. Influence of reactive oxygen species on *De novo* acquisition of resistance to bactericidal antibiotics. *Antimicrob Agents Chemother* 2018;62:e02354-17.
3. Burmeister AR, Fortier A, Roush C, Lessing AJ, Bender RG, *et al*. Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. *Proc Natl Acad Sci U S A* 2020;117:11207–11216.
4. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 2010;8:15–25.
5. Wetterstrand KA. DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) [cited 15/03/2021]; (n.d.). www.genome.gov/sequencingcostsdata
6. Le VVH, Davies IG, Moon CD, Wheeler D, Biggs PJ, *et al*. Novel 5-nitrofuran-activating reductase in *Escherichia coli*. *Antimicrob Agents Chemother* 2019;63:e00868-19.
7. Lázár V, Nagy I, Spohn R, Csörgő B, Györkei Á, *et al*. Genome-wide analysis captures the determinants of the antibiotic cross-resistance interaction network. *Nat Commun* 2014;5:4352.
8. Praski Alzrigat L, Huseby DL, Brandis G, Hughes D. Resistance/fitness trade-off is a barrier to the evolution of MarR inactivation mutants in *Escherichia coli*. *J Antimicrob Chemother* 2021;76:77–83.
9. Hill CW, Gray JA. Effects of chromosomal inversion on cell fitness in *Escherichia coli* K-12. *Genetics* 1988;119:771–778.
10. Fitzgerald SF, Lupolova N, Shaaban S, Dallman TJ, Greig D, *et al*. Genome structural variation in *Escherichia coli* O157:H7. *Microb Genom* 2021;7:11. ;
11. Guérillot R, Kostoulias X, Donovan L, Li L, Carter GP, *et al*. Unstable chromosome rearrangements in *Staphylococcus aureus* cause phenotype switching associated with persistent infections. *Proc Natl Acad Sci U S A* 2019;116:20135–20140.
12. Cui L, Neoh H, Iwamoto A, Hiramatsu K. Coordinated phenotype switching with large-scale chromosome flip-flop inversion observed in bacteria. *Proc Natl Acad Sci U S A* 2012;109:E1647-56.
13. Wardell SJT, Gauthier J, Martin LW, Potvin M, Brockway B, *et al*. Genome evolution drives transcriptomic and phenotypic adaptation in *Pseudomonas aeruginosa* during 20 years of infection. *Microb Genom* 2021;7:11.
14. Savic DJ, Nguyen SV, McCullor K, McShan WM. Biological impact of a large-scale genomic inversion that grossly disrupts the relative positions of the origin and terminus loci of the *Streptococcus pyogenes* chromosome. *J Bacteriol* 2019;201:17.
15. Shukla SK, Kislow J, Briska A, Henkhaus J, Dykes C. Optical mapping reveals a large genetic inversion between two methicillin-resistant *Staphylococcus aureus* strains. *J Bacteriol* 2009;191:5717–5723.
16. Casse F, Pascal MC, Chippaux M. Comparison between the chromosomal maps of *Escherichia coli* and *Salmonella typhimurium*. Length of the inverted segment in the trp region. *Mol Gen Genet* 1973;124:253–257.
17. Alokam S, Liu SL, Said K, Sanderson KE. Inversions over the terminus region in *Salmonella* and *Escherichia coli*: IS200s as the sites of homologous recombination inverting the chromosome of *Salmonella enterica* serovar typhi. *J Bacteriol* 2002;184:6190–6197.
18. Hayashi K, Morooka N, Yamamoto Y, Fujita K, Isono K, *et al*. Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol Syst Biol* 2006;2:0007.
19. Le VVH, Biggs PJ, Wheeler D, Davies IG, Rakonjac J. Novel mechanisms of TolC-independent decreased bile-salt susceptibility in *Escherichia coli*. *FEMS Microbiol Lett* 2020;367:10.
20. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, *et al*. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2006;2:2006.
21. Cox MP, Peterson DA, Biggs PJ. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. *BMC Bioinformatics* 2010;11:485.
22. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
23. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, *et al*. The sequence alignment/ap format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
24. R-Core-Team. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing;
25. Garrison E, Marth G. (n.d.) Haplotype-based variant detection from short-read sequencing. *arXiv 2012:arXiv:12073907 [q-bioGN*
26. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, *et al*. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* 2012;6:80–92.
27. Boža V, Perešíni P, Brejová B, Vinař T. DeepNano-blitz: a fast base caller for MinION nanopore sequencers. *Bioinformatics* 2020;36:4191–4192.
28. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 2018;34:2666–2669.
29. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, *et al*. Trycycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 2021;22:266.
30. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019;37:540–546.
31. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. *F1000Res* 2019;8:2138.
32. Li H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 2016;32:2103–2110.
33. Vaser R, Šikić M. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 2021;1:332–336.
34. Wick RR, Holt KE. Polypolish: short-read polishing of long-read bacterial genome assemblies. *Bioinformatics* 2021. DOI: 10.1101/2021.10.14.464465.
35. Zimin AV, Salzberg SL. The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol* 2020;16:e1007981.
36. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:e11147.
37. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res* 2004;14:1394–1403.
38. Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 2018;34:3094–3100.
39. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, *et al*. Integrative genomics viewer. *Nat Biotechnol* 2011;29:24–26.
40. Dashti AA, Jadaon MM, Abdulsamad AM, Dashti HM. Heat treatment of bacteria: a simple method of DNA extraction for molecular techniques. *Kuwait Med J* 2009;41:117–122.
41. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, *et al*. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 2012;13:134.
42. Wiegand I, Hilpert K, Hancock REW. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat Protoc* 2008;3:163–175.
43. Wickham H. ggplot2. In: *Ggplot2: Elegant Graphics for Data Analysis*. Cham: Springer-Verlag New York, 2016.
44. Skovgaard O, Bak M, Løbner-Olesen A, Tommerup N. Genome-wide detection of chromosomal rearrangements, indels, and mutations in circular chromosomes by short read sequencing. *Genome Res* 2011;21:1388–1393.
45. Kono N, Tomita M, Arakawa K. eRP arrangement: a strategy for assembled genomic contig rearrangement based on replication profiling in bacteria. *BMC Genomics* 2017;18:784.

46. Goldberg A, Fridman O, Ronin I, Balaban NQ. Systematic identification and quantification of phase variation in commensal and pathogenic *Escherichia coli*. *Genome Med* 2014;6:11.
47. Band VI, Hufnagel DA, Jaggavarapu S, Sherman EX, Wozniak JE, *et al*. Antibiotic combinations that exploit heteroresistance to multiple drugs effectively control infection. *Nat Microbiol* 2019;4:1627–1635.
48. Rueggeberg KG, Toba FA, Thompson MG, Campbell BR, Hay AG. A Q-like transcription factor regulates biofilm development in *Escherichia coli* by controlling expression of the DLP12 lysis cassette. *Microbiology (Reading)* 2013;159:691–700.
49. Wang X, Kim Y, Wood TK. Control and benefits of CP4-57 prophage excision in *Escherichia coli* biofilms. *ISME J* 2009;3:1164–1179.
50. Maduike NZ, Tehranchi AK, Wang JD, Kreuzer KN. Replication of the *Escherichia coli* chromosome in RNase HI-deficient cells: multiple initiation regions and fork dynamics. *Mol Microbiol* 2014;91:39–56.
51. Rudolph CJ, Upton AL, Stockum A, Nieduszynski CA, Lloyd RG. Avoiding chromosome pathology when replication forks collide. *Nature* 2013;500:608–611.
52. Veetil RT, Malhotra N, Dubey A, Seshasayee ASN. Laboratory evolution experiments help identify a predominant region of constitutive stable DNA replication initiation. *mSphere* 2020;5:e00939-19.