Pre-existing chromosomal polymorphisms in pathogenic *E. coli* potentiate the evolution of antibiotic resistance by MCR-1 plasmid acquisition

Pramod K. Jangir^{1,2#*}, Qiue Yang^{3#}, Liam P. Shaw¹, Julio Diaz Caballero¹, Lois Ogunlana¹, Rachel M. Wheatley¹, Timothy R. Walsh¹, and Craig MacLean^{1*}

¹Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK

²Present address: Biozentrum, University of Basel, Basel, Switzerland

³Fujian Provincial Key Laboratory of Soil Environmental Health and Regulation, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China

#equal contribution

*Correspondence: Craig MacLean (craig.maclea@zoo.ox.ac.uk); Pramod K. Jangir (jangirk.pramod@gmail.com; pramod.jangir@unibas.ch)

Abstract

Bacterial pathogens show high levels of standing genetic diversity, but the influence of this diversity on the evolution of antibiotic resistance remains unclear. Here we address this problem in the context of colistin, a ‘last line of defense’ antibiotic. Using experimental evolution, we show that a plasmid carrying the MCR-1 colistin resistance gene dramatically increases the ability of *E. coli* populations to evolve high-level colistin resistance by acquiring mutations in *lpxC*, an essential chromosomal gene involved in lipopolysaccharide biosynthesis. Crucially, *lpxC* mutations increase colistin resistance in the presence of the MCR-1 gene, but decrease the resistance of wild-type cells, revealing positive sign epistasis for antibiotic resistance. Analysis of public genomic datasets shows that *lpxC* polymorphisms are common in pathogenic *E. coli* carrying MCR-1, highlighting the clinical relevance of this interaction. Importantly, *lpxC* diversity is high in pathogenic *E. coli* from regions with no history of MCR-1 acquisition, suggesting that pre-existing *lpxC* polymorphisms have potentiated the evolution of high-level colistin resistance by MCR-1 acquisition. More broadly, these findings highlight the importance of standing genetic variation and plasmid/chromosomal interactions in the evolutionary dynamics of antibiotic resistance.

Key words: Antibiotic resistance, pathogen evolution, colistin, epistasis, plasmids

Introduction

Decades of research on the genetics of antibiotic resistance have shown that pathogenic bacteria evolve resistance to clinically relevant doses of antibiotic by acquiring a small number of mutations or by acquiring mobile elements carrying dedicated antibiotic resistance gene^{1,2}. An open and unresolved question in evolutionary biology and genetics is the extent to which the other genes in bacterial genomes contribute to the evolution of resistance by modifying the rate of acquisition and effects of resistance determinants^{3–10}.

Many of the bacterial pathogens where resistance poses a high risk to public health are opportunistic pathogens with high levels of chromosomal genetic diversity^{11–13}. Although these pathogens acquire resistance determinants against a backdrop of extensive genomic diversity, the importance of this diversity in shaping the evolution of resistance remains poorly understood^{10–14}. One approach to addressing this problem has been to challenge diverse collections of clinical/environmental strains with evolving resistance to a controlled antibiotic pressure using laboratory experimental evolution^{4,5,14,15}. These studies have shown that strains evolve resistance at different rates, and it has been possible to identify ‘potentiator’ genes that accelerate the evolution of resistance by epistatically increasing the effects of canonical chromosomal resistance mutations^{14,15}. For example, strains of *Staphylococcus aureus* with high levels of expression of the *norA* efflux pump rapidly evolve resistance to fluoroquinolone antibiotics because high levels of *norA* expression increase the effect of classical fluoroquinolone resistance mutations in DNA topoisomerase¹⁴. Although these studies have focused on resistance evolution by chromosomal mutations, most resistance in important pathogenic bacteria is caused by the acquisition of plasmids carrying mobile resistance genes^{1,2}. It is equally conceivable that chromosomal polymorphisms shape the evolution of resistance by altering the acquisition and fitness effects of horizontally acquired resistance genes⁸.

One important challenge of using experimental evolution to identify potentiator genes is that pathogen isolates often differ from each other by thousands of single nucleotide polymorphism (SNPs) and variation in the presence/absence of accessory genes. Previous studies have used transcriptomic methods and GWAS (genome-wide association study) to identify genes that cause variation in evolvability between strains, but identifying these causal variants is challenging using this approach^{10–14}. Given that potentiator genes interact

epistatically with resistance genes, it follows that the presence of a resistance gene should also generate selection on potentiator genes under antibiotic pressure. According to this logic, comparing genomic changes between populations that carry established resistance determinants and sensitive control populations that arise during selection for elevated antibiotic resistance should provide a simple way to search for potentiator genes that interact epistatically with classical ‘resistance’ genes.

To test this idea, we challenged populations of *E. coli* carrying an MCR-1 (mobile colistin resistance) plasmid and wild-type control populations with adapting to colistin, a last-line of defence antibiotic that is used in the treatment of infections caused by multidrug-resistant pathogens¹⁶. *E. coli* has adapted to the widespread use of colistin in agriculture by acquiring plasmids carrying the MCR-1 gene. MCR-1 is selectively advantageous in the presence of colistin^{17,18}, but it generates a colistin resistance phenotype that is of borderline clinical significance^{17,19–21}. In our experiments, MCR-1 populations consistently evolved high-level colistin resistance by acquiring mutations in *lpxC*, an essential gene involved in lipopolysaccharide (LPS) biosynthesis that interacts epistatically with MCR-1 to increase resistance. Wild-type populations, on the other hand, were unable to evolve increased colistin resistance under the strong selective regime used in our experiments. We then tested for associations between MCR-1 and *lpxC* polymorphisms in pathogenic and commensal *E. coli* using large genomic data sets to better understand the consequences of this interaction *in vivo*.

Results

Accelerated evolution of colistin resistance in MCR-1 *E. coli*

To search for genes that interact with MCR-1 to increase colistin resistance, we evolved populations of *E. coli* carrying MCR-1 encoded on an IncX4 plasmid (MCREC) and wild-type control populations under strong selection for increased colistin resistance. Our selection experiment used an ‘evolutionary ramp’ in which the concentration of colistin was doubled daily from the very low dose (1/8 minimum inhibitory concentration (MIC)) to a very high dose (16x MIC) (see Methods). Antibiotic doses were standardized relative to the MICs of the two parental strains to ensure that the strength of selection was equivalent for the two strains. The rapid increase in antibiotic concentration in the ‘evolutionary ramp’ treatment ensures that populations must either evolve increased resistance or face extinction (once

concentrations exceed the MIC of the parental strains)^{15,22}. Thus, the rate of populations extinction defines the measure of the evolvability of a strain.

Previous selection experiments have shown that mutants with low resistance tend to have high resistance evolvability, generating characteristic patterns of diminishing returns adaptation⁹ that has been reported for other microbial phenotypes^{23,24}. All else being equal, we would therefore expect that control populations (initial MIC=0.5mg/L) would have higher evolvability than the MCREC populations (initial MIC=4mg/L). In contrast to this expectation, MCREC populations showed consistent evidence of adaptation to colistin, whereas only a very small minority of control populations were able to evolve increased resistance (Figure 1; Proportional Hazard likelihood ratio = 116, $P = <0.0001$).

Figure 1. MCR-1 carrying bacterial populations (MCREC) show accelerated evolution of high-level colistin resistance. The plot shows the survival of wild-type and MCR-1 carrying populations of *E. coli* ($n = 120$ populations for *E. coli* and $n = 60$ populations for *E. coli*+MCR-1 (PN23 IncX4) over time under increasing selection for colistin resistance. MCR-1 carrying populations showed increased survival, which is indicative of higher colistin evolvability. The numbers shown at the bottom represent the colistin concentration for wild-type *E. coli* (black) and for *E. coli* with MCR-1 (orange).

Mechanistic basis of the evolution of high-level colistin resistance

The increased evolvability of MCREC populations suggests that the presence of MCR-1 opened up new pathways to evolving elevated colistin resistance that were not accessible to wild-type populations. To test this idea, we sequenced the genomes of n=21 independently evolved MCREC clones from day 6 of the experiment (4x MIC). All of the evolved clones had increased colistin resistance (mean MIC = 19.2 mg/L, range = 8-32mg/L) compared to the MCREC parental strain (MIC= 4mg/L) (Supplementary Table 1).

Figure 2. (a) List of mutations identified in colistin evolved clones. **(b)** Figure shows the mutations identified in LpxC. LpxC positions 79, 238, 242 and 246 are important for Zn binding and LpxC activity²⁵. nt - nucleotide; Δ - deletion; SNP - single nucleotide polymorphism; Indel - insertions/deletions of nucleotides

Strikingly, the majority of clones (n=11/21) carried mutations in LpxC (Figure 2a), an UDP-3-O-acyl-N-acetylglucosamine deacetylase that catalyzes the first step in lipid A biosynthesis and is important for membrane composition and asymmetry^{26,27}. *lpxC* is an essential gene²⁸, suggesting that its evolution should be highly constrained. Consistent with this idea, *lpxC* mutations were clustered in a small number of domains and both of the observed deletions kept the coding sequence in frame. Interestingly, several clones had mutations in the metal (zinc) binding domain site which is very important for *lpxC* activity²⁵ (Figure 2b). Mutations in other genes involved in LPS biosynthesis were found in evolved clones that lacked *lpxC* mutations, including *lpxA* (n=2) and *rfaY* (n=1) (Figure 2a). Interestingly, a minority of clones lacked mutations and/or structural variants (n=4). The colistin resistance of these clones did not differ suggesting that adaptive plasticity in gene expression and/or unstable genetic changes (i.e., heteroresistance²⁹) can also contribute to increased colistin resistance (Supplementary Figure 1).

Testing for interactions between LpxC and MCR-1

The presence of *lpxC* mutations in MCREC populations, suggests that the presence of MCR-1 increases the selective advantage provided by *lpxC* mutations in the presence of colistin. To test this hypothesis, we focused on measuring the interaction between MCR-1 and a *lpxC* mutation from a domain that is critical for LpxC function (i.e., I244N). The MCR-1 natural plasmid used in our experiments (i.e., IncX4:MCR-1) is an intermediate size plasmid (approx. 34 kb) that also carries genes for housekeeping and conjugation, as well as cargo genes of unknown function²¹. To examine the interaction between LpxC and MCR-1 more directly, we cloned MCR-1 into pSEVA121, a low copy number (approx. 5 copies/cell) synthetic plasmid that confers a similar level of colistin resistance (MIC=5mg/L) to the IncX4:MCR-1 plasmid used the selection experiment (MIC=4mg/L). The key advantage of the pSEVA system is that this makes it possible to measure the impact of realistic levels of MCR-1 expression while controlling for any background plasmid effects using an empty vector control.

Our approach to measure the interaction between LpxC and MCR-1 was to test for epistasis between these genes in quantitative traits that were key targets of selection in our experiment. Epistasis occurs when different genes make a non-additive contribution to phenotypes³⁰. For any quantitative trait, epistasis (e) can be calculated as $e = w_{AB} * w_{ab} - w_{Ab}w_{aB}$ where w_{AB} is the value of the trait in the wild type, ab is the double mutants, and the single

mutants are represented as Ab and aB ³¹. Negative epistasis ($e < 0$) occurs when the value of the trait is less in the double mutant than would be expected from the single mutant, and vice versa for positive epistasis ($e > 0$). Although evolutionary biologists typically focus on epistatic interactions for fitness, the same methods can be applied to calculate epistasis for any quantitative trait.

The rapid increase in colistin concentration in our experiment implies that colistin resistance is likely to have been a key target of selection. MCR-1 increased the colistin resistance of the *E. coli* parental strain from 0.5mg/L to 4mg/L, which is the clinical breakpoint MIC for colistin resistance³². The *lpxC* mutation, on the other hand, marginally decreased colistin resistance in the wild-type strain (~35% decrease in MIC) (Figure 3a). However, the combination of *lpxC* and MCR-1 increased colistin resistance to about 2x higher than the levels observed in the MCR-1 strain, revealing positive sign epistasis between *lpxC* and MCR-1 for colistin resistance (Figure 3a).

Electrostatic attractions occur between colistin, which is positively charged, and the bacterial outer membrane, which is negatively charged. MCR-1 increases resistance to colistin in part by reducing the net negative charge on the bacterial outer membrane by modifying LPS³³. Given that *lpxC* is involved in LPS biosynthesis, we reasoned that *lpxC* and MCR-1 may interact epistatically to reduce membrane surface charge. As expected, the MCR-1 gene led to a substantial reduction in cell surface charge (Figure 3b). The *lpxC* mutation led to a small reduction in net cell surface charge that was not statistically significant (Wilcoxon rank-sum test $P = 0.150$). However, the combination of the *lpxC* mutation and MCR-1 reduced the surface charge to a greater extent than the MCR-1 alone (Figure 3b). The surface charge of the *lpxC*-MCR-1 double mutant was marginally less than the additive expectation, providing evidence of weak negative epistasis for cell surface charge ($e = -0.059$, $\text{error} = 0.0355$).

Antibiotic resistance is associated with costs that modulate the selective benefit of resistance³⁴. Previous work has shown that MCR-1 expression is associated with fitness costs²¹, and it is possible that *lpxC* mutations provide a fitness advantage by compensating for the costs of MCR expression. To address this possibility, we measured the fitness of mutants in antibiotic-free culture medium. As expected, MCR-1 was associated with a small, but detectable, fitness cost (Figure 3c). *lpxC* is an essential gene and, unsurprisingly, the *lpxC* mutation was associated with a high fitness cost in a wild-type background (Figure 3c). The combination of the *lpxC* mutation and MCR-1 double mutant had lower fitness than either of

the single mutants, allowing us to rule out the possibility that *lpxC* mutations represent a compensatory adaptation for MCR-1. Interestingly, the fitness of double mutant was marginally less than we would expect under an additive model, revealing weak negative epistasis for fitness ($\varepsilon=-0.0244$, error=0.0089).

Figure 3. Effect of the *lpxC* mutation (*lpxC**), with and without MCR-1, on bacterial susceptibility to colistin resistance (a), membrane surface charge (b), and fitness (c). (a) Bacterial susceptibility to colistin was tested by measuring MIC ($n = 3$ biological replicates per strain). (b) Relative change in membrane surface charge was measured by using FITC-PLL binding assay ($n = 5$ biological replicates/strain) (see Methods). The changes were compared to wild-type (WT) control strain. (c) Bacterial fitness was measured using in-vitro competition assay where strains were competed against a tester *E. coli* strain carrying a chromosomally integrated GFP ($n = 5$ biological replicates per competition). Statistical significance was determined by pairwise comparisons using the Wilcoxon rank-sum test, and double asterisks show differences with a P value < 0.01 . Central horizontal lines represent mean values of biological replicates. WT – wild-type, EV – empty vector

Association between MCR-1 and LpxC in pathogenic *E. coli*

The large-scale use of colistin as a growth promoter in agriculture drove the spread of MCR-1 across agricultural, environmental and human-associated populations of *E. coli* in South East Asia^{18,35}. To test the importance of *lpxC* in colistin resistance, we searched for *lpxC* polymorphisms in a recently published data set of 688 MCREC isolates from China collected from a diversity of agricultural, environmental, and human-associated sources³⁶. Notably, the frequency of *lpxC* polymorphisms in pathogenic *E. coli* isolates collected from infections was higher than the level found in patients colonized with *E. coli* and in the environment (Figure 4). Colistin is used as a last line of defence antibiotic for the treatment of infections caused by multidrug-resistant pathogens¹⁶ and this association between *lpxC* and MCR-1 in pathogenic *E. coli* is worrying, as it suggests that high-level colistin resistance may be common in pathogenic MCR-1-positive isolates.

Figure 4. *mcr-1*-positive isolates from patients infected with Enterobacteriaceae have more mutations in *lpxC* compared to isolates from other sources. Data shown are from n=673 unfiltered isolates collected in 2016 by Shen et al. (2020)³⁶. (a) SNPs in *lpxC* determined by alignment of the *lpxC* sequence for each assembled isolate against the reference *lpxC* sequence from *E. coli* MG1655 (NC_000913.3:106557-107474). Mean values (+/- standard error) are shown by red points (lines). Isolates from patients infected with Enterobacteriaceae have significantly more SNPs than other sources, for example, those from patients only colonized with Enterobacteriaceae (Wilcoxon rank-sum test, $W=1150$, $P<0.001$). ***= $P<0.001$. However, this difference is associated with (b) the population structure of *E. coli*, with a greater proportion of phylogroups B2 and D in isolates from infected patients. The same plot for filtered isolates (Supplementary Figure 2) showed a similar pattern although the comparison between patients infected vs. colonized was not significant.

Although this association between MCR and *lpxC* mutations in pathogenic *E. coli* appears compelling, the population structure of *E. coli* differs between isolate sources. In particular, extraintestinal pathogenic *E. coli* usually belong to phylogroup B2 and D^{37–39}, with B2 the main lineage for the majority of bloodstream *E. coli* infections in the UK⁴⁰. Similarly, the majority of B2 isolates in the Shen dataset were from patients infected with Enterobacteriaceae (10/18). This suggested the importance of considering the population structure of *E. coli* when considering standing genetic variation in *lpxC*. It is not clear whether these *lpxC* mutations were acquired after MCR-1 (as in our experiments) or if these are pre-existing *lpxC* polymorphisms. To address this issue with a wider dataset, we analyzed *lpxC* diversity in genomic data sets of commensal, pathogenic and animal-associated *E. coli* from the UK, where MCR-1 is effectively absent^{41,42}. The diversity of *lpxC* was consistently highest in isolates from the B2 phylogroup in this MCR-1 naïve population of *E. coli* (Figure 5).

The vast majority of SNPs seen in *lpxC* were synonymous (n=6746/6783), suggesting that *lpxC* is under strong purifying selection. Only 37/1429 isolates had non-synonymous alleles. Strikingly, the majority were within B2 from the Davies infection dataset (33/37, 89%), despite B2 making up only around half of the filtered dataset (n=659, 46%). This greater proportion of isolates with nonsynonymous SNPs in B2 compared to other phylogroups was significant (Chi-squared test $p < 0.001$, X-squared=26.7; 33/659 B2 vs. 44/771 non-B2). We verified this higher diversity of *lpxC* in phylogroup B2 using complete RefSeq assemblies (Supplementary Figure 3). However, RefSeq isolates did not show a greater proportion of nonsynonymous SNPs in B2, with both having ~2% of isolates with nonsynonymous SNPs (Chi-squared test $p = 1$, X-squared < 0.001; 6/278 B2 vs. 35/1611 non-B2). It may be relevant that the Davies dataset was from bloodstream infections whereas RefSeq includes a diverse range of isolates.

The greater diversity of *lpxC* in the pathogenic B2 phylogroup across studies is intriguing. Levels of overall genetic diversity are known to differ between *E. coli* phylogroups, with phylogroup D having the highest nucleotide diversity across 112 core *E. coli* gene families in a recent study⁴³. However, phylogroup B2 was found to have similar overall nucleotide diversity to major phylogroups such as A or B1, suggesting the observation of higher diversity in *lpxC* is not simply an artefact of higher overall genetic diversity. Altogether, this analysis suggests that pathogenic populations of *E. coli* already contain higher levels of standing

genetic variation in *lpxC*, which may pre-dispose them to evolving high-level colistin resistance by MCR-1 acquisition.

Figure 5. *LpxC* diversity in main *E. coli* phylogroups. The diversity of *lpxC* within phylogroup B2 is greater across three different datasets (called Davies⁴¹, Rehab⁴², Shen³⁶) (see Methods for details). Left: boxplots of the pairwise SNPs between *lpxC* sequences of the same phylogroup within different studies. Dashed line shows the median pairwise distance (6 SNPs) between isolates of different phylogroups. Right: a midpoint-rooted neighbour-joining tree of *lpxC* diversity, with colored tips in each panel showing the location of isolates from each main phylogroup. Assembled isolates were filtered to only include high-quality assemblies (see Methods).

Conclusion

The acquisition of specialized antibiotic resistance genes through horizontal gene transfer has played a key role in the evolution of antibiotic resistance in pathogenic bacteria. One of the key insights from genomics studies has been that human pathogens are highly polymorphic^{11,12}, but the role of chromosomal genetic variation in shaping the acquisition and fitness effects of mobile resistance genes remains poorly understood^{8,44,45}. Here we show that

the evolution of high-level colistin resistance is driven by epistatic interactions between the MCR-1 mobile colistin resistance and *lpxC*, a chromosomal gene involved in LPS biosynthesis (Figure 1-3). *lpxC* mutations are more common in pathogenic MCR-positive *E. coli* (MCREC) than commensal MCREC, highlighting the clinical relevance of this interaction (Figure 4). Crucially, *lpxC* polymorphisms are over-represented in pathogenic lineages for *E. coli* from populations with no history of MCR-1 acquisition, suggesting that pre-existing polymorphisms in *lpxC* potentiated the evolution of high-level colistin resistance by MCR-1 acquisition (Figure 5). The underlying causes of high *lpxC* diversity in pathogenic *E. coli* are unclear, and we hope that future work will shed light on the impact of *lpxC* polymorphisms on *E. coli* virulence and infection.

One of the most surprising features of our results is that approximately 20% of clones that were recovered from MCR-1-positive populations growing at 4xMIC (16mg/L) did not have any detectable mutations or structural variants, a phenomenon that we have also reported in *P. aeruginosa*⁴⁶. There are a number of examples of adaptation to colistin through the acquisition of unstable resistance mutations that rapidly revert when selection for resistance is relaxed (i.e., heteroresistance²⁹). Importantly, we cultured clones for DNA extraction on a medium containing a low dose of colistin (1 mg/L, ¼ MIC) that may not have generated effective selection for the maintenance of unstable resistance determinants. Alternatively, it is possible that these clones were sampled from populations that adapted to colistin phenotypically, for example, due to the growth of a sub-set of cells with altered expression of genes involved in LPS biosynthesis. Given the importance of colistin as a ‘last line of defence’ antibiotic, an important goal for future work will be to investigate the mechanisms that underpin reversible colistin resistance and how these interact with MCR-1.

There is growing evidence that standing genetic variation (i.e., pre-existing variation) in pathogen species has played an important role in the evolution of antibiotic resistance. For example, at least two of the major lineages of methicillin-resistant *Staphylococcus aureus* (MRSA) originated well before the clinical introduction of methicillin^{47,48}. Our study builds on previous work^{4,5,14,15} by demonstrating that epistatic interactions between mobile resistance genes and chromosomal variants shape the evolution of resistance. We speculate that plasmid-chromosome epistasis is likely to be particularly important for highly mobile resistance genes, such as MCR-1^{17,35,49} and carbapenemases⁵⁰ that are rapidly transferred

between divergent bacterial strains and species by conjugation. Our findings suggest that understanding interactions between mobile resistance genes and background genetic variation could help to develop methods for predicting the spread of resistance in pathogen populations, which is an important goal for evolution and microbiology.

Acknowledgements

This work was supported by grants from the Wellcome Trust (106918/Z/15Z, C.M.) and the Medical Research Council (MR/S013768/1, T.W. and C.M.). L.O. was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) doctoral training partnership (BB/M011224/1). Y.Q.E. is supported by the National Natural Science Foundation of China (31200150) and Science Foundation of Fujian province (2021J01116). L.P.S. is a Sir Henry Wellcome Postdoctoral Fellow funded by Wellcome (grant 220422/Z/20/Z). We thank Csaba Pal lab (Biological Research Centre, Szeged, Hungary) for providing the pORTMAGE plasmid. We also thank Ákos Nyerges (Harvard University) for the helpful discussion on pORTMAGE.

Author contributions

C.M., P.K.J., and Y.Q.E. designed research; P.K.J., Y.Q.E., L.O., and R.W. performed experiments. L.P.S. performed *E. coli* public genome data collection and analysis. J.D.C. performed genome sequencing data analysis. C.M., P.K.J., L.P.S., T.R.W., and J.D.C. interpreted the data. C.M. and P.K.J. wrote the manuscript with input from all co-authors.

Competing interests

The authors declare no competing interests.

Data availability

All data generated or analyzed during this study are included in this article and its Supplementary Information. All sequencing data will be available soon on the NCBI GenBank.

Methods

Bacterial strains and growth condition

All the experiments were carried out in *E. coli* J53. Experiments were conducted in Luria-Bertani (LB) medium (Sigma-Aldrich).

pSEVA:MCR-1 vector construction

A synthetic MCR-1 plasmid was constructed by cloning *mcr-1* gene into pSEVA121 plasmid⁵¹. The *mcr-1* gene along with its natural promoter was PCR-amplified from the natural PN16 (IncI2) plasmid using Q5® High-Fidelity DNA Polymerase (New England BioLabs). The amplified and purified *mcr-1* fragment was assembled together with PCR-amplified pSEVA121 backbone using NEBuilder® HiFi DNA Assembly Master Mix according to the manufacturer's instructions. The assembled plasmid was then transformed into *E. coli* J53 using the standard electroporation method. Cells were recovered in 1 ml SOC medium, followed by 1 h incubation at 37°C. Different dilutions of transformant mixture were made and were plated onto Petri dishes containing LB agar supplemented with 100 µg/ml ampicillin. The culture plates were incubated at 37°C for overnight. PCR and sequence verification by Sanger sequencing were performed to ensure the presence of the correctly assembled recombinant plasmid. A full list of the primers used is given in Supplementary Table 2.

Experimental evolution

The experimental evolution experiment was carried out using our previously established 'evolutionary ramp' method^{15,22}. Briefly, wild-type *E. coli* (control) and *E. coli* with PN23 IncX4:MCR-1 plasmid were grown on LB agar medium supplemented with appropriate antibiotic at 37°C for overnight. To ensure reproducibility, we had 120 and 60 independent replicates for *E. coli* control and *E. coli*-MCR-1 strain (MCREC), respectively. Parallel cultures were propagated in 96-well microtiter plates where the inner 60 wells were used for bacterial cultures and the outer wells for blank measurement and contamination controls. We used checkerboard layout on the plate to control for potential cross-contaminations. The plates were incubated for 22 h and then overnight grown cultures were diluted 1:400-fold into fresh LB medium plates with increasing concentration of colistin. The concentration of colistin was doubled daily from the very low dose (1/8 MIC) to a very high dose (16x MIC). At the end of each transfer, the optical density was measured, and the experiment was carried out until more than >90% of cultures have gone extinct (i.e., OD₅₉₅ less than 0.05). In total, eight

transfers were performed in LB medium with colistin. The evolved *E. coli*-MCR-1 and control populations were preserved at -80°C in 15% glycerol solution. We tested for a difference between the survival of control populations (n=120) and plasmid carrying populations (n=60) using Cox's proportional hazard model. Each population was considered extinct on the day when OD fell below our 0.05 threshold.

DNA extraction, whole-genome sequencing and data analysis

To identify the polymorphisms associated with colistin resistance evolution, we selected 21 evolved populations from day 6 (4X MIC). Bacterial DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen) on a QIAcube automated system (Qiagen). The concentration of each genomic DNA sample was determined using a QuantiFluor dsDNA System (Promega) following the manufacturer's protocol. The purified genomic DNA samples were subjected to Illumina HiSeq 4000 platform at the Oxford Genomics Centre (Wellcome Centre for Human Genetics, University of Oxford, UK). The ancestral isolate was sequenced on an Oxford nanopore MinION platform using a FLO-MIN106 flow-cell and the SQK-LSK109 Ligation Sequencing kit.

Raw Illumina reads were quality controlled with Trimmomatic v. 0.39⁵² using the ILLUMINACLIP (2:30:10) and SLIDING WINDOW (4:15) parameters. Raw nanopore reads were basecalled using guppy v. 0.0.0+7969d57 and reads were demultiplexed using qcat v. 1.1.0 (<https://github.com/nanoporetech/qcat>). Resulting sequencing reads were assembled using unicycler v. 0.4.8⁵³, which used SAMtools v. 1.9⁵⁴, pilon v. 1.23⁵⁵, and bowtie2 v. 2.3.5.1⁵⁶, in hybrid mode with respective Illumina reads. The resulting assembly was annotated using Prokka v. 1.14.0⁵⁷.

Small variants were identified using breseq v. 0.34.0⁵⁸ in the polymorphism prediction mode, which implemented bowtie2 v. 2.3.5.1. Copy number variants were estimated with the default parameters of the package CNOGpro v. 1.1⁵⁹ in R⁶⁰. The plasmid copy number was calculated by dividing the average sequencing coverage of each plasmid by the average sequencing coverage of the chromosome using the SAMtools depth algorithm (Supplementary Figure 4-6).

Construction of I244N *lpxC* mutant (*lpxC)**

A previously described pORTMAGE method⁶¹ was applied to introduce I244N *lpxC* mutation into the wild-type *E. coli* J53. The mutant was constructed via synthetic ssDNA-mediated

recombineering using the pORTMAGE3 plasmid. The MODEST tool was used to design 90 nucleotides long ssDNA oligos that had complementary sequences to the replicating lagging strand with a minimized secondary structure ($\geq -12 \text{ kcal mol}^{-1}$)⁶². The oligos were ordered from Integrated DNA Technologies (Coralville, IA, USA). For the recombineering, pORTMAGE3 vector was transformed into *E. coli* J53 electrocompetent cells using standard electroporation method. *E. coli* carrying pORTMAGE3 vector was further used for recombineering. pORTMAGE3 was induced at 42°C for 15 min to allow for efficient mutation incorporation and to avoid off-target mutagenesis. 1 µl of the 100 µM mutation-carrying oligo was transformed into induced electrocompetent cells. Following the electroporation, cells were recovered in 5 mL Terrific-Broth (TB) media (24 g yeast extract, 12 g tryptone, 9.4 g K₂HPO₄, and 2 g KH₂PO₄ per 1 L of water) and incubated at 30 °C for 60 min. After the incubation period, 5 mL of LB medium was added, and cells were further incubated at 30°C overnight. Appropriate dilutions of the cultures were then plated onto LB agar plates to form individual colonies and incubated at 30°C overnight. Individual colonies were screened for the mutation using Sanger capillary-sequencing. Finally, the pORTMAGE3 plasmid was cured from the sequence-verified colony by growing the cells at 42°C for overnight in antibiotic-free LB medium. A list of oligos used for the mutation construction is given in Supplementary Table 2.

Determination of minimum inhibitory concentration (MIC)

MICs were determined using a standard serial broth dilution method with a minor modification that we previously optimized to capture small differences in MIC⁶³. Specifically, smaller colistin concentration steps were used (typically 1.5-fold). 10-steps serial dilution of colistin was prepared in fresh LB medium in 96-well microtiter plates. Three wells contained only LB medium to monitor the bacterial growth in the absence of colistin. Bacterial strains were grown in LB medium supplemented with appropriate antibiotic (100 µg/ml ampicillin for *E. coli* pSEVA:MCR-1) at 30°C for overnight. Following overnight incubation, approximately 5x10⁵ cells were inoculated into the each well of the 96-well microtiter plate. Three independent replicates for each strain and the corresponding control were used. The top and bottom row in the 96-well plate were filled with LB medium to obtain the background OD value of the growth medium. Plates were incubated at 30 °C with continuous shaking at 250 rpm. After 20–24 h of incubation, OD₆₀₀ values were measured in a microplate reader

(Biotek Synergy 2). After background subtraction, MIC was defined as the lowest concentration of colistin where the OD₆₀₀ < 0.05.

Membrane surface charge measurement

To measure bacterial membrane surface charge, we carried out a fluorescein isothiocyanate-labelled poly-L-lysine (FITC-PLL) (Sigma) binding assay^{63,64}. FITC-PLL is a polycationic molecule that binds to an anionic lipid membrane in a charge-dependent manner and is used to investigate the interaction between cationic peptides and charged lipid bilayer membranes. The assay was performed as previously described⁶³. Briefly, bacterial cells were grown overnight in LB medium, and then centrifuged, and washed twice with 1X PBS buffer (pH 7.4). The washed bacterial cells were re-suspended in 1× PBS buffer to a final OD₆₀₀ of 0.1. A freshly prepared FITC-PLL solution was added to the bacterial suspension at a final concentration of 6.5 µg/ml. The suspension was incubated at room temperature for 10 min, and pelleted by centrifugation. The remaining amount of FITC-PLL in the supernatant was determined fluorometrically (excitation at 500 nm and emission at 530 nm) with or without bacterial exposure. The quantity of bound molecules was calculated from the difference between these values. A lower binding of FITC-PLL indicates a less net negative surface charge of the outer bacterial membrane.

Analysis of *E. coli* genomes datasets

We downloaded three short-read datasets of *E. coli* genomes to investigate *lpxC* diversity: *mcr-1*-positive *E. coli* from China from a range of sources including human infection isolates, collected 2016-2018 (n=688; Shen dataset)³⁶; *mcr-1*-negative *E. coli* bloodstream infection isolates from Oxfordshire, UK, collected 2013-2015 (n=981 short-read datasets from n=976 isolates; Davies dataset)⁴¹; and *mcr-1*-negative *E. coli* from UK, mainly from livestock, collected 2017 (n=502; REHAB dataset)⁴².

Scripts and data for analysis, including assemblies, are available on figshare (doi: 10.6084/m9.figshare.18591347). In brief, genomes for the Shen and Davies datasets were assembled from short reads downloaded from the NCBI SRA. These reads were filtered with Trimmomatic v0.39 with default parameters⁵² and assembled with Spades v3.15.3⁶⁵. Genomes from the REHAB dataset were downloaded from previously assembled genomes from hybrid assembly. We screened genomes for a full-length *lpxC* using blastn against *lpxC* from *E. coli* MG1655 (NZ_LR881938.1_cds_WP_000595482.1_97; identical to *lpxC* in strain

J53; both phylogroup A). We discarded genomes without a full-length *lpxC* match (n=5), those where *lpxC* had >12 SNPs compared to the reference suggesting poor assembly (n=20, all from Shen dataset) leaving n=2146 isolates to analyse. Details of these isolates including accessions, phylogroup, and SNPs in *lpxC* are given in Supplementary Table 3. We assigned phylogroup with ClermonTyper and analysed the n=2127 isolates which were in the main phylogroups (A, B1, B2, C, D, E, F, G). To visualise *lpxC* diversity we plotted a midpoint-rooted neighbour-joining tree including the reference. After obtaining initial results similar to Figure 5 (not shown) we subsequently filtered isolates to include only higher-quality assemblies (L90<100, <999 contigs) to ensure that assembly quality was not driving the patterns, resulting in n=1429 isolates. Numbers given in Results refer to this filtered dataset.

To reproduce our finding of higher *lpxC* diversity in phylogroup B2 on another dataset, we downloaded all complete *E. coli* assemblies on NCBI RefSeq (as of 13 Jan 2022, n=1903). Of these, n=1887 (>99%) were from the main phylogroups and had a full-length *lpxC* with <13 SNPs against the reference *lpxC*.

References

1. MacLean, R. C. & San Millan, A. The evolution of antibiotic resistance. *Science* (80-.). **365**, 1082–1083 (2019).
2. Partridge, S. R., Kwong, S. M., Firth, N. & Jensen, S. O. Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin. Microbiol. Rev.* **31**, (2018).
3. Lukačišinová, M., Fernando, B. & Bollenbach, T. Highly parallel lab evolution reveals that epistasis can curb the evolution of antibiotic resistance. *Nat. Commun.* **11**, (2020).
4. Vogwill, T., Kojadinovic, M., Furió, V. & MacLean, R. C. Testing the Role of Genetic Background in Parallel Evolution Using the Comparative Experimental Evolution of Antibiotic Resistance. *Mol. Biol. Evol.* **31**, 3314–3323 (2014).
5. Bethke, J. H. *et al.* Environmental and genetic determinants of plasmid mobility in pathogenic *Escherichia coli*. *Sci. Adv.* **6**, (2020).
6. Card, K. J., LaBar, T., Gomez, J. B. & Lenski, R. E. Historical contingency in the evolution of antibiotic resistance after decades of relaxed selection. *PLOS Biol.* **17**, e3000397 (2019).
7. Vogwill, T., Kojadinovic, M. & MacLean, R. C. Epistasis between antibiotic resistance mutations and genetic background shape the fitness effect of resistance across species of *Pseudomonas*. *Proc. R. Soc. B Biol. Sci.* **283**, 20160151 (2016).
8. Bottery, M. J., Wood, A. J. & Brockhurst, M. A. Adaptive modulation of antibiotic resistance through intragenomic coevolution. *Nat. Ecol. Evol.* **1**, 1364–1369 (2017).
9. MacLean, R. C., Perron, G. G. & Gardner, A. Diminishing Returns From Beneficial Mutations and Pervasive Epistasis Shape the Fitness Landscape for Rifampicin Resistance in *Pseudomonas aeruginosa*. *Genetics* **186**, 1345–1354 (2010).
10. Torres Ortiz, A. *et al.* Genomic signatures of pre-resistance in Mycobacterium tuberculosis. *Nat. Commun.* **12**, 7312 (2021).
11. Baker, S., Thomson, N., Weill, F.-X. & Holt, K. E. Genomic insights into the emergence and spread of antimicrobial-resistant bacterial pathogens. *Science* (80-.). **360**, 733–738 (2018).
12. Didelot, X., Bowden, R., Wilson, D. J., Peto, T. E. A. & Crook, D. W. Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **13**, 601–612 (2012).

13. Smith, J. M., Feil, E. J. & Smith, N. H. Population structure and evolutionary dynamics of pathogenic bacteria. *BioEssays* **22**, 1115–1122 (2000).
14. Papkou, A., Hedge, J., Kapel, N., Young, B. & MacLean, R. C. Efflux pump activity potentiates the evolution of antibiotic resistance across *S. aureus* isolates. *Nat. Commun.* **11**, 3970 (2020).
15. Gifford, D. R. *et al.* Identifying and exploiting genes that potentiate the evolution of antibiotic resistance. *Nat. Ecol. Evol.* **2**, (2018).
16. Li, J. *et al.* Colistin: the re-emerging antibiotic for multidrug-resistant Gram-negative bacterial infections. *Lancet Infect. Dis.* **6**, 589–601 (2006).
17. Liu, Y.-Y. *et al.* Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* **16**, 161–168 (2016).
18. Sun, J., Zhang, H., Liu, Y.-H. & Feng, Y. Towards Understanding MCR-like Colistin Resistance. *Trends Microbiol.* **26**, 794–808 (2018).
19. Bardet, L., Le Page, S., Leangapichart, T. & Rolain, J.-M. LBJMR medium: a new polyvalent culture medium for isolating and selecting vancomycin and colistin-resistant bacteria. *BMC Microbiol.* **17**, 220 (2017).
20. Doumith, M. *et al.* Detection of the plasmid-mediated *mcr-1* gene conferring colistin resistance in human and food isolates of *Salmonella enterica* and *Escherichia coli* in England and Wales. *J. Antimicrob. Chemother.* **71**, 2300–2305 (2016).
21. Yang, Q. *et al.* Balancing *mcr-1* expression and bacterial survival is a delicate equilibrium between essential cellular defence mechanisms. *Nat. Commun.* **8**, 2054 (2017).
22. San Millan, A., Escudero, J. A., Gifford, D. R., Mazel, D. & MacLean, R. C. Multicopy plasmids potentiate the evolution of antibiotic resistance in bacteria. *Nat. Ecol. Evol.* **1**, (2017).
23. Kryazhimskiy, S., Rice, D. P., Jerison, E. R. & Desai, M. M. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science (80-.).* **344**, 1519–1522 (2014).
24. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-Term Dynamics of Adaptation in Asexual Populations. *Science (80-.).* **342**, 1364–1367 (2013).
25. Jackman, J. E., Raetz, C. R. H. & Fierke, C. A. Site-Directed Mutagenesis of the

568 Bacterial Metalloamidase UDP-(3- O -acyl)- N -acetylglucosamine Deacetylase (LpxC).
569 Identification of the Zinc Binding Site. *Biochemistry* **40**, 514–523 (2001).

570 26. Guest, R. L., Rutherford, S. T. & Silhavy, T. J. Border Control: Regulating LPS
571 Biogenesis. *Trends Microbiol.* **29**, 334–345 (2021).

572 27. Fivenson, E. M. & Bernhardt, T. G. An Essential Membrane Protein Modulates the
573 Proteolysis of LpxC to Control Lipopolysaccharide Synthesis in *Escherichia coli*. *MBio*
574 **11**, (2020).

575 28. Beall, B. & Lutkenhaus, J. Sequence analysis, transcriptional organization, and
576 insertional mutagenesis of the *envA* gene of *Escherichia coli*. *J. Bacteriol.* **169**, 5408–
577 5415 (1987).

578 29. Andersson, D. I., Nicoloff, H. & Hjort, K. Mechanisms and clinical relevance of
579 bacterial heteroresistance. *Nat. Rev. Microbiol.* **17**, 479–496 (2019).

580 30. Durão, P., Balbontín, R. & Gordo, I. Evolutionary Mechanisms Shaping the
581 Maintenance of Antibiotic Resistance. *Trends Microbiol.* **26**, 677–691 (2018).

582 31. Trindade, S. *et al.* Positive Epistasis Drives the Acquisition of Multidrug Resistance.
583 *PLoS Genet.* **5**, e1000578 (2009).

584 32. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables
585 for interpretation of MICs and zone diameters. *Version 12.0* (2022).

586 33. Sabnis, A. *et al.* Colistin kills bacteria by targeting lipopolysaccharide in the
587 cytoplasmic membrane. *Elife* **10**, (2021).

588 34. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: is it possible to
589 reverse resistance? *Nat. Rev. Microbiol.* **8**, 260–271 (2010).

590 35. Wang, R. *et al.* The global distribution and spread of the mobilized colistin resistance
591 gene *mcr-1*. *Nat. Commun.* (2018) doi:10.1038/s41467-018-03205-z.

592 36. Shen, C. *et al.* Dynamics of *mcr-1* prevalence and *mcr-1*-positive *Escherichia coli* after
593 the cessation of colistin use as a feed additive for animals in China: a prospective
594 cross-sectional and whole genome sequencing-based molecular epidemiological
595 study. *The Lancet Microbe* **1**, (2020).

596 37. Picard, B. *et al.* The Link between Phylogeny and Virulence in *Escherichia coli*
597 Extraintestinal Infection. *Infect. Immun.* **67**, 546–553 (1999).

598 38. Johnson, J. R., Delavari, P., Kuskowski, M. & Stell, A. L. Phylogenetic Distribution of
599 Extraintestinal Virulence-Associated Traits in *Escherichia coli*. *J. Infect. Dis.* **183**, 78–88

(2001).

39. Moulin-Schouleir, M. *et al.* Extraintestinal Pathogenic *Escherichia coli* Strains of Avian and Human Origin: Link between Phylogenetic Relationships and Common Virulence Patterns. *J. Clin. Microbiol.* **45**, 3366–3376 (2007).
40. Kallonen, T. *et al.* Systematic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131. *Genome Res.* **27**, 1437–1449 (2017).
41. Davies, T. J. *et al.* Reconciling the Potentially Irreconcilable? Genotypic and Phenotypic Amoxicillin-Clavulanate Resistance in *Escherichia coli*. *Antimicrob. Agents Chemother.* **64**, (2020).
42. Shaw, L. P. *et al.* Niche and local geography shape the pangenome of wastewater- and livestock-associated Enterobacteriaceae. *Sci. Adv.* **7**, (2021).
43. Touchon, M. *et al.* Phylogenetic background and habitat drive the genetic diversification of *Escherichia coli*. *PLOS Genet.* **16**, e1008866 (2020).
44. McNally, A. *et al.* Combined Analysis of Variation in Core, Accessory and Regulatory Genome Regions Provides a Super-Resolution View into the Evolution of Bacterial Populations. *PLOS Genet.* **12**, e1006280 (2016).
45. San Millan, A., Toll-Riera, M., Qi, Q. & MacLean, R. C. Interactions between horizontally acquired genes create a fitness cost in *Pseudomonas aeruginosa*. *Nat. Commun.* **6**, 6845 (2015).
46. Kapel, N., Caballero, J. D. & MacLean, C. Localized hypermutation drives the evolution of unstable colistin resistance in small populations of *Pseudomonas aeruginosa*. *bioRxiv Prepr.* (2021).
47. Harkins, C. P. *et al.* Methicillin-resistant *Staphylococcus aureus* emerged long before the introduction of methicillin into clinical practice. *Genome Biol.* **18**, 130 (2017).
48. Gill, J. L., Hedge, J., Wilson, D. J. & MacLean, R. C. Evolutionary Processes Driving the Rise and Fall of *Staphylococcus aureus* ST239, a Dominant Hybrid Pathogen. *MBio* **12**, (2021).
49. Wang, Y. *et al.* Comprehensive resistome analysis reveals the prevalence of NDM and MCR-1 in Chinese poultry production. *Nat. Microbiol.* **2**, 16260 (2017).
50. León-Sampedro, R. *et al.* Pervasive transmission of a carbapenem resistance plasmid in the gut microbiota of hospitalized patients. *Nat. Microbiol.* **6**, 606–616 (2021).

51. Silva-Rocha, R. *et al.* The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res.* **41**, (2013).
52. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
53. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput. Biol.* **13**, e1005595 (2017).
54. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
55. Walker, B. J. *et al.* Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. *PLoS One* **9**, e112963 (2014).
56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
57. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
58. Deatherage, D. E. & Barrick, J. E. Identification of Mutations in Laboratory-Evolved Microbes from Next-Generation Sequencing Data Using breseq. in 165–188 (2014). doi:10.1007/978-1-4939-0554-6_12.
59. Brynildsrud, O., Snipen, L.-G. & Bohlin, J. CNOGpro: detection and quantification of CNVs in prokaryotic whole-genome sequencing data. *Bioinformatics* **31**, 1708–1715 (2015).
60. Dalgaard, P. R. D. C. T. R: A language and environment for statistical computing. (2010).
61. Nyerges, Á. *et al.* A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. *Proc. Natl. Acad. Sci.* **113**, 2502–2507 (2016).
62. Bonde, M. T. *et al.* MODEST: a web-based design tool for oligonucleotide-mediated genome engineering and recombineering. *Nucleic Acids Res.* **42**, W408–W415 (2014).
63. Kintses, B. *et al.* Chemical-genetic profiling reveals limited cross-resistance between antimicrobial peptides with different modes of action. *Nat. Commun.* **10**, 5731 (2019).
64. Rossetti, F. F. *et al.* Interaction of Poly(L-Lysine)-g-Poly(Ethylene Glycol) with

- 664 Supported Phospholipid Bilayers. *Biophys. J.* **87**, (2004).
- 665 65. Bankevich, A. *et al.* SPAdes: A New Genome Assembly Algorithm and Its Applications
666 to Single-Cell Sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
667
668
669

Supplementary information

Pre-existing chromosomal polymorphisms in pathogenic *E. coli* potentiate the evolution of antibiotic resistance by MCR-1 plasmid acquisition

Pramod K. Jangir^{1,2#*}, Qiue Yang^{3#}, Liam P. Shaw¹, Julio Diaz Caballero¹, Lois Ogunlana¹, Rachel M. Wheatley¹, Timothy R. Walsh¹, and Craig MacLean^{1*}

¹Department of Zoology, University of Oxford, Oxford, OX1 3PS, UK

²Present address: Biozentrum, University of Basel, Basel, Switzerland

³Fujian Provincial Key Laboratory of Soil Environmental Health and Regulation, College of Resources and Environment, Fujian Agriculture and Forestry University, Fuzhou 350002, China

#equal contribution

*Correspondence: Craig MacLean (craig.maclean@zoo.ox.ac.uk); Pramod K. Jangir (jangirk.pramod@gmail.com; pramod.jangir@unibas.ch)

Supplementary Figure 1. Figure comparing the colistin resistance of clones with mutations in *lpxA/C*, other targets, and no mutations (i.e., none). There was no significant difference between the groups (Wilcoxon rank-sum test). Each clone represents one data point.

Supplementary Figure 2. A version of Figure 4 showing SNPs in *lpxC* in the Shen dataset with filtering criteria for isolate assemblies applied (see Methods).

Supplementary Figure 3. Analysing NCBI RefSeq complete assemblies shows that B2 isolates have more SNPs relative to the reference *lpxC* than other phylogroups, reproducing the finding from the three assembled datasets (Figure 5).

Supplementary Figure 4. Sequencing reads of each population were mapped to the hybrid assembly of reference d0-5-1. The sequence of plasmid PN23 was assembled in a single sequence, and the sequence of plasmid p193 was distributed in three contigs. Plasmid copy number for each population was calculated by dividing the average depth of coverage of each plasmid by the depth of coverage of the chromosome.

Supplementary Figure 5. Sequencing coverage (y axis) of each coding region (black dots) in the chromosome of each population. X-axis represent the position of each CDS in the respective replicon. Coverage was normalized given its GC content and calculated using CNOGpro⁵⁹ and the d0-5-1 reference chromosome sequence.

Supplementary Figure 6. Sequencing coverage (y axis) of each coding region (black dot) in the PN23 plasmid of each population. X-axis represent the position of each CDS in the respective replicon. Coverage was normalized given its GC content and calculated using CNOGpro⁵⁹ and the d0-5-1 reference PN23 plasmid sequence.