Supplementary material

Bacterial culture and isolate selection

Putative *Enterobacteriaceae* isolates were selected for further analysis on the basis of typical colonial morphology - namely moist-looking colonies which were cream, white or yellow in color and with a regular edge.

The sweeps of growth from feces and individual isolates harvested were re-cultured from frozen stocks by scraping the stock with an inoculating loop and plating onto the agar. The sweeps of growth feces were cultured on CBA without antibiotic; the isolates on MacConkey supplemented with cefotaxime at 2mg/L.

16s rDNA sequencing

Performed as described in [1].

mcr-1 PCR and *bla*$_{\text{CTX-M}}$ PCR/allele sequencing

Isolates were screened for *mcr-1* by PCR using the primers CLR5-F (5′-CGGTCAGTCCGTTTGTTC-3′) and CLR5-R (5′-CTTGGTCGGTCTGTA GGG-3′) and amplicons were subsequently sequenced[2]. *bla*$_{\text{CTX-M}}$ and *bla*$_{\text{CTX-M}}$-group alleles were screened for using the primers available in the online supplementary information in[3].

DNA extraction process

DNA was extracted for sequencing by adding cetyl-trimethyl-ammonium bromide (CTAB) buffer to the sample, vortexing and incubating at 37°C for one hour. Chloroform was added to cooled samples; these were then centrifuged at 12,000rpm

(10mins, 25°C), and the top, aqueous layer transferred to a new Eppendorf. Isopropanol (-20°C) and sodium acetate (3M, 10%) were then mixed with the sample, which was stored at (-20°C) overnight. Centrifugation was repeated as above; this time, the top layer was discarded and 750µL ethanol added to the sediment. The sample was centrifuged at 12,000rpm (5mins, 25°C); the top, aqueous layer was discarded, and the DNA was dried and re-dissolved in TE buffer.

Library preparation for Illumina sequencing

For paired-end libraries, 1.5µg of each DNA extract was fragmented by ultra-sonication (Covaris Inc, Woburn, MA, USA), and correct fragment size confirmed by gel electrophoresis and photometric scanning. Fragmented DNA was combined with End Repair Mix (Illumina), and incubated at 20°C for 30 min. End-repaired DNA was purified with the QIAquick PCR Purification Kit (Qiagen), added to the A-Tailing Mix, and incubated at 37°C for 30 min. Adapters were then ligated to the 3'-adenylated fragments by incubating with Adapter and Ligation mixes at 20°C for 15 min. Adapter-ligated DNA was selected by running a 2% agarose gel to recover target fragments, and then purified with the QIAquick Gel Extraction kit (Qiagen). Fragments were amplified with the PCR Primer Cocktail and PCR Master Mix, and recovered and purified using the QIAquick PCR Purification Kit (Qiagen), as above.

For mate-pair libraries, 20µg DNA was fragmented by ultra-sonication, and the fragments end-repaired with the biotin dNTPs mix, by incubating at 20°C for 30 min. End-repaired DNA was purified with the QIAquick PCR Purification Kit(Qiagen), and target fragments selected by running a 0.6% agarose gel. Ligation Mix was added, and the mixture incubated at 16°C overnight. DNA fragments were

circularized, and non-circularized DNA removed by digestion. Circular DNA was again fragmented (Covaris), and biotinylated purified using Dynabeads M-280 streptation (Invitrogen). End-repair Mix was added, and the mixture incubated at 20°C for 30 min. End-repaired DNA was purified with streptavidin beads, A-Tailing Mix was added, and the mixture incubated at 37°C for 30 min. Adenylated DNA, Adapter and Ligation Mix, were then incubated at 20°C for 15 min. Adapter-ligated DNA was purified, and amplified by PCR using the PCR Primer Cocktail and PCR Master Mix. 400-600bp PCR products were selected by running a 2% agarose gel and purified as above.

Phylogenetic methods

The phylogenetic model implemented in IQtree was a generalized time reversible nucleotide substitution model and a gamma distribution allowing for substitution rate variation among sites (GTR+G setting), using a starting tree inferred by maximum parsimony.

ClonalFrameML settings were default, with kappa estimated from the IQtree result output (2.75543).

Plasmid contig circularization and Bandage analysis: Assembly visualization and inference of plasmid structure for non-circularizable *mcr-1* plasmid contigs

Plasmid contigs were considered circular, closed structures if BLASTn matches of contig ends overlapped by ≥50bp, demonstrated 100% sequence identity, and were unique within the contig. From the assembly graph, we inferred the likely connectivity of any contigs associated with the *mcr-1*-harboring contig as far as

possible based on the contig linkages and individual contig coverage statistics. High contig coverage (>2-fold) relative to the coverage for contigs known to be chromosomal was used as a proxy marker for multi-copy/plasmid associated sequences. For these high coverage contigs (nodes) the connectivity of any nodes associated with the *mcr-1*-harboring contig was determined based on the consistency of coverage between those nodes. If there were two pairs of nodes (I1, I2, O1, O2) going in and out of another high coverage node (N), with clear differences between the coverage values between nodes (e.g. I1 and O2 have similar coverage c1, and I2 and O1 has similar coverage c2) then a connection was made between the pair of in/out nodes (i.e. in this example I1-N-O2 and I2-N-O1). The longest possible path containing *mcr-1* was then extracted using the "Output/Specify exact path for copy/save" function in Bandage.

Iterative sequential regression (ISR)

To investigate the prevalence of fecal samples harboring *mcr-1*-positive isolates, and *mcr-1*-positive/cefotaxime-resistant isolates, respectively, we modeled counts of each entity per month using Poisson regression with the log of the number of fecal samples per month as the offset. To identify changes in trends over time we used ISR[4]. In brief, we started by comparing a one-slope model to a 2-slopes model (i.e. incorporating a breakpoint) between April 2011 and March 2012 with October 2011 set as the initial breakpoint. If the 2-slopes model was a better fit (determined by a Bayesian Information Criterion lower by ≥3.84 [the critical value to detect a significance level of 0.05 with a χ^2 test and one degree of freedom]) we fixed the breakpoint and looked at the next one-year period, otherwise, we increased the initial endpoint by one month (e.g. in the first instance to April 2012) and compared the

model with one slope to the models with 2 slopes with either October 2011 or November 2011 as a breakpoint. We then iterated over this the process, with October 2015 as the last possible breakpoint (6 months of data required either side of the breakpoint to allow detection).

Figure creation

Additional graphical representations were undertaken using Geneious[5] and BioPython's GenomeDiagram module[6].

Supplementary Table/Figure Legends

Table S1. Summary of epidemiological (demographic and laboratory) and genomic features of study isolates.

Figure S1. Alignment of the novel plasmid structure identified in this study with a non-*mcr-1* IncP plasmid, pHNFP671.

Supplementary references

1. Ye H, Li Y, Li Z, Gao R, Zhang H, Wen R, Gao GF, Hu Q, Feng Y: **Diversified *mcr-1*-Harbouring Plasmid Reservoirs Confer Resistance to Colistin in Human Gut Microbiota.** *MBio* 2016, **7:**e00177.
2. Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, 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* 2016, **16:**161-168.

3. Tian GB, Huang YM, Fang ZL, Qing Y, Zhang XF, Huang X: **CTX-M-137, a hybrid of CTX-M-14-like and CTX-M-15-like beta-lactamases identified in an *Escherichia coli* clinical isolate.** *J Antimicrob Chemother* 2014, **69:**2081-2085.

4. Schlackow I, Walker AS, Dingle K, Griffiths D, Oakley S, Finney J, Vaughan A, Gill MJ, Crook DW, Peto TE, Wyllie DH: **Surveillance of infection severity: a registry study of laboratory diagnosed *Clostridium difficile*.** *PLoS Med* 2012, **9:**e1001279.

5. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al: **Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.** *Bioinformatics* 2012, **28:**1647-1649.

6. Pritchard L, White JA, Birch PR, Toth IK: **GenomeDiagram: a python package for the visualization of large-scale genomic data.** *Bioinformatics* 2006, **22:**616-617.