1 **High rates of human faecal carriage of *mcr-1*-positive multi-drug resistant isolates** 2 **emerge in China in association with successful plasmid families**

3

4 Lan-Lan Zhong^{1,2*}, Hang TT Phan^{3*}, Xi Huang^{1,2}, Karina Doris-Vihta³, Anna E
5 Sheppard³, Kun-Jiao Zeng^{1,2}, Hong-Yu Li⁴, Xue-Fei Zhang^{1,2}, Sandip Patil^{1,2}, Yan-Fen
6 Zhang^{1,2}, Cong Shen^{1,2}, Derrick W Crook³, A Sarah Walker³, Yong Xing^{1,2}, Qian-yi
7 Chen^{1,2}, Jia-lin Lin^{1,2}, Lian-Qiang Feng^{1,2}, Yohei Doi⁵, Nicole Stoesser^{3†}, Guo-Bao
8 Tian^{1,2†}

9

10 **Author affiliations:**

11 ¹ Department of Immunology, Zhongshan School of Medicine, Sun Yat-sen University,
12 74 Zhongshan 2nd Road, Guangzhou 510080, China

13 ² Key Laboratory of Tropical Diseases Control (Sun Yat-sen University), Ministry of
14 Education, Guangzhou 510080, China

15 ³ Modernising Medical Microbiology, Nuffield Department of Medicine, University of
16 Oxford, Oxford, United Kingdom

17 ⁴ Department of Clinical Laboratory, Sun Yat-sen Memorial Hospital, Sun Yat-sen
18 University, Guangzhou 510080, China

19 ⁵ University of Pittsburgh Medical Centre, Pittsburgh, Pennsylvania 15261, USA

20

21 * These authors contributed equally to this article

22 † These senior authors contributed equally to this article

23 **Short running title:** Human faecal carriage of *mcr-I* in China

24

25 **Corresponding author:**

26 Dr Nicole Stoesser, MBBS, D.Phil.

27 Mailing address: Department of Microbiology (Research), John Radcliffe Hospital Level

28 7, Headley Way, Headington, OX3 9DU, UK.

29 E-mail: nicole.stoesser@ndm.ox.ac.uk

30 Tel: +44 (0)1865 220856

31 Fax: +44 (0)1865 764192

32

33 **Alternate corresponding author:**

34 Dr Guo-Bao Tian, Ph.D.

35 Mailing address: Department of Immunology, Zhongshan School of Medicine, Sun Yat-

36 sen University, 74 Zhongshan 2nd Road, Guangzhou 510080, China.

37 E-mail: tiangb@mail.sysu.edu.cn

38 Fax/Phone: +86 (0)20 87335387

39

40 **Word count:** Synopsis 249/250, main text 3500 [50 references]

41 Table: 1, Figures: 5

42 Supplementary material, Supplementary table: 1, Supplementary figures: 2

43 **SYNOPSIS**

44 **Background**

45 *mcr-I*-mediated transmissible colistin resistance in Enterobacteriaceae is concerning,
46 given colistin is frequently used as a treatment of last resort in multidrug-resistant
47 Enterobacteriaceae infections. Reported rates of human *mcr-I* gastrointestinal carriage
48 have historically been low.

49 **Objectives**

50 To identify trends in human gastrointestinal carriage of *mcr-I* positive and *mcr-I*-
51 positive/cefotaxime-resistant Enterobacteriaceae in Guangzhou, China, 2011-2016, and
52 investigate the genetic contexts of *mcr-I* in a subset of *mcr-I*-positive/cefotaxime-
53 resistant strains using whole genome sequencing (WGS).

54 **Methods**

55 Of 8,022 faecal samples collected, 497 (6.2%) were *mcr-I*- positive, and 182 (2.3%) *mcr*-
56 *I*-positive/cefotaxime-resistant. Trends in carriage were assessed using iterative
57 sequential regression. A subset of *mcr-I*-positive isolates was sequenced (Illumina), and
58 genetic contexts of *mcr-I* were characterised.

59 **Results**

60 We observed marked increases in *mcr-I* (now ~30% prevalence) and more recent (since
61 January 2014) increases in *mcr-I*-positive/third-generation cephalosporin-resistant
62 Enterobacteriaceae human colonisation (p<0.001). Sub-cultured *mcr-I*-positive/third-
63 generation cephalosporin-resistant isolates were commonly multi-drug resistant.

64

WGS of 50 *mcr-1*/third-generation cephalosporin-resistant isolates (49 *Escherichia coli*; 1 *Klebsiella pneumoniae*) demonstrated bacterial strain diversity (39 *E. coli* sequence types); *mcr-1* in association with common plasmid backbones (IncI, IncHI2/HI2A, IncX4) and sometimes in multiple plasmids; frequent *mcr-1* chromosomal integration; and loss of the *mcr-1*-associated insertion sequence IS*AplI* in some plasmids. Significant sequence similarity with published *mcr-1* plasmid sequences was consistent with spread amongst pig, chicken and human reservoirs.

Conclusions

The high positivity rate (~10%) of *mcr-1* in multidrug-resistant *E. coli* colonising humans is a clinical threat; the diverse genetic mechanisms (strains/plasmids/insertion sequences) associated with *mcr-1* have likely contributed to its dissemination, and will facilitate its persistence.

77 MAIN TEXT

78 INTRODUCTION

79 The polymyxins, including colistin, are antibiotics of last resort for managing multidrug-
80 resistant Gram-negative infections, including *Pseudomonas aeruginosa*, *Acinetobacter*
81 spp. and Enterobacteriaceae. Acquired resistance in these organisms has historically
82 largely been due to: mutation-based modifications to cell wall lipopolysaccharides;
83 utilization of efflux pumps; and capsule formation[1]. Transmissible colistin resistance,
84 encoded by *mcr-1*, was recently identified in *Escherichia coli* and *Klebsiella pneumoniae*
85 isolates from hospitalized humans, animals (pigs) and raw meat (pigs and chicken) in
86 China[2], with higher rates in animal samples (~19% versus ~1% in humans).

87
88 Subsequently, *mcr-1*-harboring strains have been variably identified in humans, animals
89 and raw meat in Europe, Asia, North America, South America and Africa (e.g.[3-13].
90 These strains have predominantly been *E. coli*[14] or *Salmonella* spp.[3, 5], with up to
91 20% carriage prevalence in swine and poultry[6, 11], and *mcr-1*-positive isolates from
92 chickens as early as the 1980s in China and 2007 in France[6, 15]. Prevalence in humans
93 remains low[2, 4, 7, 10], and mostly restricted to hospitalized patients[16]. However *mcr-*
94 *1* can be carried in the healthy human gut[17].

95
96 The association of *mcr-1* with other broad-spectrum resistance mechanisms, such as
97 extended-spectrum β -lactamases (ESBLs) and/or carbapenemases[18-23], has major
98 clinical implications. As polymyxins are last-line therapies for carbapenem-resistant
99 infections, widespread, joint dissemination of these mechanisms would represent a major

problem for case management. The identification of *mcr-I* in multiple plasmid types, including IncI2[2, 23], IncHI2[20], IncX4[12, 23], IncP[24] and IncF[21] replicons, is consistent with multiple *mcr-I* mobilization events, potentially facilitating the association of *mcr-I* with other resistance mechanisms, thereby creating multidrug-resistant bacterial hosts.

105

Here we investigate human faecal carriage of *mcr-I* in Guangzhou, China, over 5 years, using whole-genome sequencing (WGS) to characterize the genetic structures involved in a subset of isolates and identify associations with other antimicrobial resistance genes.

109

110 MATERIALS AND METHODS

8,022 consecutive faecal samples, initially collected to detect ESBL-producing Enterobacteriaceae, were submitted to the laboratory at Sun-Yat Sen Memorial Hospital 1st April 2011-31st March 2016. Samples came from three hospitals (Sun-Yat Sen Memorial Hospital, Guangdong General Hospital, and Qingyuan People's Hospital) in Guangzhou, the capital city of Guangdong province, serving a population of ~15 million over ~10,000 km² in southeast China. All samples were obtained continuously, except for January 2012, February 2013 and February 2014 (holiday months of the Spring Festival; staff shortages). Ethical approval for the study was given by Sun Yat-Sen University; individual consent for the use of faecal samples was obtained for all patients.

120

Faecal samples were plated onto Columbia blood agar (CBA) without antibiotic within 2 hours of collection and incubated for 18-24 hours at 37°C. Subsequently, up to 10

colonies with the appearance of Enterobacteriaceae were sub-cultured to MacConkey agar (Thermo Fisher, Beijing, China) supplemented with cefotaxime (2 mg/L), and species identification performed by 16S rDNA sequencing (Supplementary material). All cefotaxime-resistant Enterobacteriaceae isolates were stored in lysogeny broth (LB) with 30% glycerol at -80°C. Sweeps from the original CBA plates were also stored in LB with 30% glycerol at -80°C.

All frozen sweeps of cultured feces (n=8,022) and cefotaxime-resistant Enterobacteriaceae isolates (n=20,332) were subsequently re-cultured from frozen stock and screened for *mcr-1* by PCR (Supplementary material). Cefotaxime-resistant isolates were screened for *bla*_{CTX-M}, and alleles determined by sequencing (Supplementary material). Species identification for *mcr-1* positive isolates was performed by API20E (bioMérieux, Marcy l'Etoile, France). Minimum inhibitory concentrations (MICs) were determined for all *mcr-1*-positive isolates by agar dilution and interpreted using the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoints (version 6.0), as well as the Clinical and Laboratory Standards Institute (document M100-S25).

All cefotaxime-resistant, *mcr-1*-positive Enterobacteriaceae of distinct species and with differing MIC profiles to May 2015 (n=45), and a random subset from June 2015-March 2016 (n=44/142 [31%]) were selected for WGS. Each stored, frozen isolate was inoculated into 50 mls LB, cultured overnight, centrifuged at 3000rpm, and transferred to the Beijing Genomics Institute (BGI) for DNA extraction and WGS. DNA was extracted

146 using the cetyl-trimethyl-ammonium bromide (CTAB)/chloroform method
147 (Supplementary material). Multi-locus sequence type was available for a subset (25/187)
148 using a standard approach (protocols at <http://mlst.warwick.ac.uk/mlst/dbs/Ecoli>).
149
150 DNA extracts were sequenced on the Illumina HiSeq 4000 platform, using both paired-
151 end (150bp reads, ~350bp insert) and mate-pair (50bp reads, ~6kb insert) approaches
152 (n=69 isolates) or paired-end reads only (n=20 isolates; Table S1). Libraries were
153 prepared using standardized protocols incorporating fragmentation by ultra-sonication,
154 end repair, adaptor ligation, and PCR amplification (Supplementary material).
155
156 Preliminary species identification for isolates was derived from WGS using Kraken[25];
157 read-data were then mapped to a species-specific reference[26]. Hybrid *de novo*
158 assemblies of paired-end and mate-pair data, or paired-end data alone, were generated
159 using SPAdes[27] version 3.6 with the “--careful” option, a set of automatically
160 determined k-mer values (21, 33, 55, 77), and by removing contigs <500bp or with k-mer
161 coverage <1. *In silico* MLST was determined using BLASTn and publicly available
162 databases for relevant species (<http://mlst.warwick.ac.uk/mlst/dbs/Ecoli>,
163 <http://bigsdh.web.pasteur.fr/klebsiella/klebsiella.html>). Resistance genes were identified
164 from *de novo* assemblies using an in-house, curated database of resistance genes[28] and
165 BLASTn/mapping-based identification (scripts available at:
166 <https://github.com/hangphan/resisType>). Contigs were annotated using PROKKA[29].
167 Contigs containing *mcr-I* were defined as “plasmid” if they contained annotations
168 consistent with plasmid-associated loci and no obvious chromosomal loci, or

169 “chromosomal” if >75% of annotations (excluding hypothetical proteins) were consistent
170 with a chromosomal location.

171

172 The integrity of assemblies containing *mcr-I* in chromosomal locations was assessed
173 using the REAPR tool[30]. We excluded any sequences with assembly sizes ≥ 6.5 Mb;
174 mixed calls from *in silico* MLST or species identification; and mismatches between
175 available laboratory-based and WGS *in silico* MLST.

176

177 Phylogenies were reconstructed using IQtree[31] with a generalized time reversible
178 nucleotide substitution model and a gamma distribution allowing for substitution rate
179 variation among sites (GTR+G setting), using a starting tree inferred by maximum
180 parsimony. Branch lengths were corrected for recombination using ClonalFrameML
181 (default settings, kappa estimated from IQtree=2.75543)[32]. Insertion sequences (IS)
182 were downloaded from ISFinder (<https://www-is.biotoul.fr>); sequence assemblies were
183 queried with BLASTn (presence requiring >95% sequence identity over >90% of the
184 reference sequence length) against this database.

185

186 Circularization of *mcr-I*-harboring plasmid contigs was confirmed using Bandage[33].
187 For single *mcr-I*-harboring plasmid contigs which could not be fully circularized, we also
188 used Bandage to visualize the sequencing assembly graph generated by SPAdes and
189 manually resolved the most likely *mcr-I* plasmid structures based on node (contig)
190 linkage and contig coverage (Supplementary material).

191

Iterative sequential regression (ISR) in R was used to characterize trends in faecal *mcr-I* positivity (Supplementary material). The phylogeny was represented in the interactive tree of life viewer (iTOL v3, <http://itol.embl.de>).

Links to raw sequencing reads for the isolates included in analyses are available through the NCBI BioProject accession number: PRJNA354216 (Table S1).

RESULTS

Of 8,022 samples collected during the 5-year study, 497 (6.2%) were *mcr-I*-positive, and 182 (2.3%) *mcr-I*-positive and cefotaxime-resistant (Fig.S1). The proportion of both *mcr-I*-positive and *mcr-I*-positive/cefotaxime-resistant faecal samples increased significantly during the study ($p < 0.001$). For *mcr-I*-positive/cefotaxime-resistant samples, this was driven by a specific increase after January 2014 ($p < 0.001$, Fig.1; 95% CI for estimated date of trend change: 01/April/2013-01/Nov/2014).

From faecal samples harbouring *mcr-I*-positive/cefotaxime-resistant Enterobacteriaceae, 187 distinct Enterobacteriaceae isolates from 179 individuals were identified (*E. coli*, $n=173$; *K. pneumoniae*, $n=13$; *Enterobacter cloacae*, $n=1$). 90 (50%) of these 179 individuals were male, median age 52 years (inter-quartile range [IQR: 39-69 years]). 144/179 (80%) individuals were hospital in-patients at the time of sampling (median length of stay 14 days [range: 3-258 days; IQR: 7-21 days]).

214 The 89 sequenced isolates comprised *E. coli* (n=79), *K. pneumoniae* (n=9), *Enterobacter*
215 *cloacae* (n=1), based on 16S rDNA sequencing. Of these, however, 30/45 consecutive
216 isolates pre-May 2015 and 9/44 randomly selected isolates post-June 2015 failed
217 sequencing quality control metrics (Table S1), leaving 50 sequences for analysis (49 *E.*
218 *coli*, 1 *K. pneumoniae*). For *E. coli*, these represented 39 sequence types (STs)(Fig.2),
219 ST156 being the most common (n=4), with most isolates (n=33) representing singleton
220 STs. Other global disease-causing lineages were also identified, including ST131, ST155
221 and ST405[34], and for *K. pneumoniae*, ST15. One pair of isolates was separated by 0
222 SNVs (SYSU0077/SYSU0078, 7 days apart), and one pair by 3 SNVs
223 (SYSU0041/SYSU0049, 28 days apart), representing likely direct/indirect transmissions
224 between patients 146 and 150, and 95 and 104, respectively. All other strains were >1126
225 SNVs apart.

226

227 A novel *mcr-I* gene variant was identified with a G3A mutation, leading to loss of the
228 first methionine in the translated protein (n=1; SYSU0052). A separate isolate had *mcr-I*
229 disrupted by an IS1294 element, previously described downstream of *mcr-I* (n=1;
230 SYSU0039)[8]. The former did not affect colistin resistance phenotype (MIC colistin=8
231 mg/L, polymixin B=16 mg/L); the latter, as expected, was colistin-susceptible (MIC
232 colistin<0.25 mg/L, polymixin B=1 mg/L).

233

234 128/187 (68%) isolates contained *bla*_{CTX-M} by PCR/sequencing (Table S1). Rates of co-
235 resistance across antimicrobial classes in *mcr-I*-positive/cefotaxime-resistant isolates
236 were high (Table 1), with only carbapenems, nitrofurantoin and tigecycline

demonstrating susceptibility rates $\geq 80\%$. Resistance was similar in sequenced isolates (Table 1). Of these, 40/50 (80%) harboured *bla*_{CTX-M}, including: (i) group 9-like alleles: 14(n=15), 27(2), 65(5), 110(1), 130(1); (ii) group 1-like alleles: 1-like(n=1), 11(1), 55/55-like(16), 136(1); and (iii) hybrid alleles: 64(2), 123(2). Eight (16%) isolates had 2 *bla*_{CTX-M} variants. In three isolates (carrying *bla*_{CTX-M-14,27/55,55/123}), a *bla*_{CTX-M} variant was located on the same plasmid contig as *mcr-1*. The presence of *bla*_{CMY}(2, 42-like [N90T]) variants explained cefotaxime resistance in two isolates. For seven (14%) isolates, no cefotaxime resistance mechanism could be identified (Table S1). Multiple other resistance genes identified included: (i) *bla*_{TEM-1} (n=37), *bla*_{TEM-135}(2), *bla*_{TEM-116}(1), *bla*_{TEM-181}(1), *bla*_{TEM-191-like}(1[K227E]), *bla*_{TEM-30-like}(1[S241R,A148P]); (ii) *bla*_{OXA-1}(2) and/or *bla*_{OXA-10}(8), *bla*_{OXA-10-like}(1[truncated, S14*]); (iii) *fosA3* (n=23); (iv) *qepA* (n=1) or *qnrS1* (n=12); *rmtB*(2) or *armA*(1); and (v) *aac(6')-Ib-cr* (6) (Table S1).

The genetic context of *mcr-1* was similarly diverse, supporting the potential for *mcr-1* to have spread widely, and in combination with other resistance genes. In 41/50 (82%) isolates *mcr-1* was located on plasmid-associated contigs, in 4 (8%) on chromosomal contigs, in 2 (4%) likely on chromosomal contigs, and in 3 (6%) the location was unclear due to assembly/annotation limitations. *mcr-1* copy number was estimated at up to 3.54 (median 0.97, inter-quartile range: 0.79-1.39); likely due to its presence either on multi-copy plasmids; as multiple copies within the same plasmid structure (SYSU0077 [IncH]; SYSU0093 [IncI2]); or on different plasmids within the same isolate (SYSU0072 [IncH and IncX4]; SYSU0220 [novel plasmid described below, IncI2-like plasmid backbone]).

260 Insertion sequences may contribute to plasmid-plasmid and plasmid-chromosome
261 rearrangements and gene mobilization. We found 99 different IS types in the 50
262 sequenced isolates, with median 19.5 (range: 9-30) IS types per isolate. The commonest
263 were IS3 (46/50 isolates), IS186B, IS2 (42/50 each), and IS150 (39/50). IS*AplI*, most
264 frequently associated with *mcr-I* mobilization to date, was found in 21/50 (42%) isolates.
265
266 In 19/41 (46%) plasmid-associated isolates *mcr-I* was co-located on the same contig as
267 an IncI2 replicon (n=17), or plausibly on an IncI2 plasmid based on assembly
268 visualization (n=2). Eleven of these 19 structures could be circularized, likely
269 representing complete plasmid structures; this included the earliest sequenced plasmid in
270 this study, and the earliest sequenced *mcr-I* plasmid, to our knowledge (isolated May
271 2011). In these 11 plasmids, the backbone, *mcr-I* location and wider gene synteny were
272 largely preserved, consistent with ancestral acquisition of *mcr-I* and subsequent
273 evolution of the plasmid structure by mutation/recombination (e.g. *repA* and the transfer
274 region, Fig.3) and acquisition/loss of smaller mobile genetic elements (MGEs). The last
275 included variable presence of ISS*en6*, a *Salmonella* spp.-associated MGE (SYSU0011,
276 SYSU0050, SYS0093); ISE*cpI*-associated acquisition of *bla*_{CTX-M-55} (SYSU0060,
277 SYSU0045); and variability in the shufflon region adjacent to the *rci* gene in the C-
278 terminal region of the *pilV* gene).
279
280 These 11 IncI2 plasmids were also genetically highly similar to three previously
281 described plasmids: SZ02 (Accession number: KU761326 [Fig.3], isolated from an
282 outpatient's blood culture, August 2015, Suzhou, China, ~1,400km away[23]); the

283 *bla*_{CTX-M-55}-harboring pA31-12 (Accession number: KX034083.1, chicken, August 2012,
284 Guangzhou), and pHNSHP45 (Accession number: KP347127.1 [Fig.3], pig, July 2013,
285 Shanghai), consistent with the dissemination of this plasmid in human cases/carriers, pigs
286 and chickens across China.

287
288 In relation to *mcr-I*, we observed variable presence of *IS**AplI* (n=5/11 [45%] study
289 isolates, reference plasmids: pHNSHP45, pA31-2). When *IS**AplI* was present, however,
290 it was consistently located 186bp upstream of *mcr-I* between *mcr-I* and *nikB*. The most
291 likely hypothesis is a historical introduction of *IS**AplI* and *mcr-I* into the IncI plasmid
292 backbone, with subsequent loss of *IS**AplI*. We also observed highly genetically related
293 plasmid backbones in two different *E. coli* host strains (ST156, 354; SYSU0060,
294 SYSU0062) isolated from the same patient on the same day, potentially consistent with
295 within-host transfer, as well as highly genetically related plasmids within different *E. coli*
296 host strains and human individuals across periods of time consistent with direct
297 transmission/acquisition from common sources (SYSU0019, SYSU0007).

298
299 In 14 (34%) isolates *mcr-I* co-localized to a contig with either IncHI2A or IncHI2+HI2A
300 (n=8), or was plausibly associated with an IncHI2 plasmid based on assembly
301 visualization. These plasmids are typically large, and we were unable to reconstruct any
302 complete IncHI2A plasmids from the sequencing data. Four contigs were short
303 (<3,000bp); the rest are represented in Fig.4, alongside two closed reference *mcr-I*
304 IncHI2 plasmids, pSA26 (Accession number: KU743384, bacteraemia, Saudi Arabia),
305 and pHNSHP45-2 (Accession number: KU341381, pig faeces isolate, Shanghai, China).

306 This alignment demonstrates a relatively homogenous plasmid backbone circulating
307 within the study population between 2014-early 2016, likely derived from an ancestral
308 plasmid structure similar to the reference plasmids, with some rearrangement/inversion
309 events, particularly around a region enriched for resistance genes and transposable
310 elements (Fig.4). Within the study isolates, sampled over 3.5 years, we observed
311 apparently frequent IS/transposon-associated indel events, including of *ISApII*, which
312 was either at the 5' end of *mcr-I* (e.g. SYSU0003), flanking it on both sides (e.g.
313 SYSU0014), or absent (e.g. SYSU0026). As *mcr-I* was located in the same wider genetic
314 context in 9/10 study plasmid contigs, this likely represents a single *ISApII-mcr-I-ISApII*
315 acquisition and subsequent loss of *ISApII* elements[35]. Alternatively it could represent a
316 “hotspot” for multiple insertion events of this complex. We also observed a *mcr-I*
317 duplication event within an otherwise identical contiguous plasmid sequence in two
318 isolates taken 7 days apart from different patients (both *E. coli* ST1684; SYSU0077,
319 SYSU0s078); and an inversion of *ISApII-mcr-I* within the plasmid backbone
320 (SYSU0002), all consistent with high rates of plasticity involving the *ISApII-mcr-I* and
321 IncHI2/HI2A structures. Other variation in the IncHI2+HI2A backbone in study isolates
322 included: (i) single nucleotide variants (SNVs), (ii) small indels (<20bp), (iii) clustered
323 SNVs suggestive of local recombination events, and (iv) gain and loss of several insertion
324 sequences.

325

326 In seven (15%) isolates *mcr-I* co-localized to (n=6) or was plausibly associated with
327 (n=1) an X4-harboring contig; in four the contigs could be fully circularized, likely
328 representing complete plasmid structures (Fig.5). These sequences had a highly

329 conserved, syntenic plasmid backbone, with limited nucleotide and indel variation,
330 including to two reference IncX4-*mcr-I* plasmids, SZ04 (inpatient peritoneal fluid,
331 Suzhou, China) and pAF48 (urine, Johannesburg, South Africa). IS*AplI* was absent in all
332 of these plasmids; the only consistently associated IS was an IS26 or highly similar
333 IS15DIV element 4.7kb upstream of *mcr-I*.

334

335 In 11 (27%) isolates no co-located or plausibly associated plasmid replicon could be
336 identified, and in one isolate (16EY645) we identified a new *mcr-I* harbouring plasmid
337 structure, p16EY645_1, of 48,944bp (Fig.S2). 94% of p16EY645_1 shared 99%
338 sequence identity with a larger IncP plasmid pHNFP671 (82,807bp, accession number:
339 KP324830.1), isolated from pig faeces obtained prior to 2014 in Guangzhou, China (the
340 study region); but p16EY645_1 itself has no typeable replicon.

341

342 Chromosomal *mcr-I* integration was confirmed in four *E. coli* STs (457, 1114, 1684,
343 2936), and plausibly occurred in two others (101, 2705), consistent with at least four-six
344 independent chromosomal integration events in diverse *E. coli* strains (12% sequenced *E.*
345 *coli* isolates), providing capacity for widespread dissemination at the organism level. In
346 three of these isolates (SYSU0021, SYSU0219, SYSU0112), IS*AplI* flanked *mcr-I* on
347 both sides (with AT-AT, AC-CC as the composite transposon target site sequences
348 [TSS], or indeterminate TSS due to contig breaks). In one isolate *mcr-I* was downstream
349 of a single IS*AplI* unit (SYSU0225); in one flanked by a previously un-described
350 ISE*c23*/IS*30*-like/IS*1294* structure (SYSU0051); and in one case inserted within a much
351 larger integrative element harbouring the *iee* gene encoding for an IS excision enhancer,

which promotes excision of members of the IS3, IS1, and IS30 families, including
IS*AplI*[36](SYSU0047).

DISCUSSION

Widespread transmissible colistin resistance is of major concern in the current era of
multidrug-resistant Gram-negative infections, as colistin is commonly used to treat
infections caused by these organisms, despite its nephrotoxicity, and issues with
determining susceptibility and appropriate dosing regimens[37, 38]. In this, the largest
study of faecal carriage and WGS of *mcr-1* isolates to date, we observed alarming
sequential increases in human faecal carriage of *mcr-1*-positive and cefotaxime-resistant,
third-generation cephalosporin-resistant *E. coli* over 5 years. We also found high genetic
plasticity and multidrug resistance of MGEs harbouring *mcr-1* in these regional data from
Guangzhou, China, that may explain some of the dramatic recent increases in human
faecal carriage over the last two years.

IncI plasmids are narrow host-range plasmids[39], commonly isolated from *E. coli* and
Salmonella spp., and implicated in the spread of the ESBL gene *bla*_{CTX-M} amongst
Enterobacteriaceae, particularly in China[40, 41]. CTX-M-55 has been found in *E. coli* in
animals in China, and CTX-M-55/55-like variants were seen with *mcr-1* in IncI2
plasmids in two different *E. coli* strains (ST156, ST117) in this study. Previously, it has
been postulated that an IncI2 plasmid backbone acquired a 3,080bp *ISEcpI*-*bla*_{CTX-M-15}
complex (GAAAA TSS) from an IncA/C plasmid, with a mutation resulting in
conversion to *bla*_{CTX-M-55} (V77A) within this structure[42]. The same signature sequence

was observed in our isolates, as well as another recently sequenced IncI2 plasmid harbouring *bla*_{CTX-M-55} from a chicken *E. coli* isolate in Guangzhou, China (August 2012)[43], consistent with the spread of this plasmid across humans and several animal species. *bla*_{CTX-M-64} was also observed on an IncI *mcr-I* plasmid in this study (SYSU0115, ST155); this allele has previously been postulated to have arisen from recombination between *bla*_{CTX-M-14} and *bla*_{CTX-M-15} on IncI plasmid backbones within food animals in China[42].

IncHI2/HI2A plasmids are typically large (>250kb)[44], multidrug-resistant plasmids that have been associated with a range of antibiotic and metal resistance genes in *Salmonella* spp. and *E. coli* isolated from humans and food-producing animals[45]. Similar to the IncI plasmids, our genetic analyses suggest an initial *mcr-I* acquisition event, in this case by direct transposition of a composite IS*AplI* transposon, and subsequent loss of IS*AplI* units and rearrangements of the plasmid backbone.

In at least four (8%) sequenced isolates, *mcr-I* had been integrated into the chromosome, a phenomenon observed in at least two other studies (*E. coli* ST156, Beijing; IS*AplI* composite transposon, TG-TG TSS[46]; *E. coli* ST410, Germany; IS*AplI* signatures consistent with composite transposition and then loss of downstream IS*AplI*, CA-CA TSS[47]). Our data suggest two additional *mcr-I* chromosomal integration events - one in association with an integrative element promoting IS activity, and one in association with ISEc23/IS30-like/IS1294 element, which has likely played a role in the mobilization of *bla*_{CMY-2} in *Salmonella* spp. and *E. coli*[48]. The multifarious means by which

398 chromosomal integration of *mcr-1* appears to occur is of major concern, as its presence in
399 the chromosome may facilitate more stable inheritance of this gene within multiple *E.*
400 *coli* strains.

401
402 IS*Apl1* is an IS30-like element derived from *Actinobacillus pleuropneumoniae*, a Gram-
403 negative organism of the Pasteurellaceae family, known to cause pleuropneumonia in
404 pigs[49], a highly infectious disease frequently warranting herd-based antimicrobial
405 treatment. IS30-like elements form circular intermediates during transposition, and have
406 been shown *in vitro* to result in 2bp target site repeats (TG, TT, AA, GG) on direct
407 transposition[49]. The variable extent of diversity observed within the plasmid families
408 here (IncHI2/HI2A and IncI > IncX4) may reflect different evolutionary rates for these
409 plasmid families, or may be consistent with earlier acquisition of IS*Apl1-mcr-1*
410 composite transposons by IncHI2/HI2A and IncI2, most likely within poultry and pig
411 farms in which regular colistin exposure represents a major selection pressure[2]. The
412 acquisition of *mcr-1* by IncX4 plasmids appears to be unrelated to the presence of
413 IS*Apl1*, and possibly associated with a rearrangement event involving IS26/26-like
414 structures (e.g. IS15DIV).

415
416 This study has several limitations. Firstly, although we observed significant increases in
417 faecal carriage of *mcr-1*-harboring isolates, we were unable to assess the possible risk
418 factors associated with increasing incidence. Secondly, our sampling strategy targeted
419 limited numbers of colistin-resistant isolates within faecal samples, and the diversity of *E.*
420 *coli* and likely other Enterobacteriaceae strains in the gastrointestinal tract is

substantial[50]. Our results could therefore underestimate the genetic diversity of strains and MGEs harbouring *mcr-1*, particularly within individual hosts. Finally, due to resource limitations we were unable to sequence all *mcr-1*-positive isolates, or to re-sequence those representing mixtures on sequencing. We were also unable to undertake any long-read sequencing, increasingly the method of choice in assembling plasmids.

Despite these limitations, we have demonstrated that human faecal carriage of *mcr-1* positive *E. coli* has increased dramatically in Guangzhou over the last two years, reaching similar proportions (20-30%) to animals (pigs, chickens) over the preceding 3-4 years in the same region of China[2]. Our genetic analyses suggest the rapid emergence of several major plasmid vectors of *mcr-1* within numerous multidrug-resistant *E. coli* strains carried by humans, and highlight the significant degree of plasticity in these plasmid vectors harbouring *mcr-1* over short periods of time.

ACKNOWLEDGEMENTS

We would like to acknowledge the work of the Modernizing Medical Microbiology (MMM) sequencing pipeline team (Nicholas Sanderson) and informatics team (Ian Szwajca, Trien Do, Dona Foster) - collectively represented as the Modernizing Medical Microbiology Informatics Group.

FUNDING

This work was supported by: the National Natural Science Foundation of China (grant 81471988); Guangdong Natural Science Foundation (S2013010015810); The 111 Project

444 (B13037, B12003); Program of Science and Technology New Star of Guangzhou
 445 (2014J2200038); and Fundamental Research Funds for the Central Universities
 446 (16ykzd09).

447

448 Additional funding support was provided through a University of Oxford/National
 449 Institutes of Health Research (NIHR) Academic Clinical Lectureship to NS and the
 450 NIHR Oxford Biomedical Research Centre. DWC and TEAP are NIHR senior
 451 investigators.

452

453 **TRANSPARENCY DECLARATIONS**

454 No conflicts of interest to declare.

REFERENCES

1. Olaitan AO, Morand S, Rolain JM: **Mechanisms of polymyxin resistance: acquired and intrinsic resistance in bacteria.** *Front Microbiol* 2014, **5**:643.
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. Quesada A, Ugarte-Ruiz M, Iglesias MR, Porrero MC, Martinez R, Florez-Cuadrado D, Campos MJ, Garcia M, Piriz S, Saez JL, Dominguez L: **Detection of plasmid mediated colistin resistance (MCR-1) in Escherichia coli and Salmonella enterica isolated from poultry and swine in Spain.** *Res Vet Sci* 2016, **105**:134-135.
4. Prim N, Rivera A, Rodriguez-Navarro J, Espanol M, Turbau M, Coll P, Mirelis B: **Detection of mcr-1 colistin resistance gene in polyclonal Escherichia coli isolates in Barcelona, Spain, 2012 to 2015.** *Euro Surveill* 2016, **21**.
5. Webb HE, Granier SA, Marault M, Millemann Y, den Bakker HC, Nightingale KK, Bugarel M, Ison SA, Scott HM, Loneragan GH: **Dissemination of the mcr-1 colistin resistance gene.** *Lancet Infect Dis* 2016, **16**:144-145.
6. Perrin-Guyomard A, Bruneau M, Houee P, Deleurme K, Legrandois P, Poirier C, Soumet C, Sanders P: **Prevalence of mcr-1 in commensal Escherichia coli from French livestock, 2007 to 2014.** *Euro Surveill* 2016, **21**.
7. Kluytmans-van den Bergh MF, Huizinga P, Bonten MJ, Bos M, De Bruyne K, Friedrich AW, Rossen JW, Savelkoul PH, Kluytmans JA: **Presence of mcr-1-**

- 478 **positive Enterobacteriaceae in retail chicken meat but not in humans in the**
479 **Netherlands since 2009. *Euro Surveill* 2016, 21.**
- 480 8. Stoesser N, Mathers AJ, Moore CE, Day NP, Crook DW: **Colistin resistance**
481 **gene mcr-1 and pHNSHP45 plasmid in human isolates of Escherichia coli**
482 **and Klebsiella pneumoniae. *Lancet Infect Dis* 2016, 16:285-286.**
- 483 9. Olaitan AO, Chabou S, Okdah L, Morand S, Rolain JM: **Dissemination of the**
484 **mcr-1 colistin resistance gene. *Lancet Infect Dis* 2016, 16:147.**
- 485 10. Kuo SC, Huang WC, Wang HY, Shiau YR, Cheng MF, Lauderdale TL: **Colistin**
486 **resistance gene mcr-1 in Escherichia coli isolates from humans and retail**
487 **meats, Taiwan. *J Antimicrob Chemother* 2016.**
- 488 11. Nguyen NT, Nguyen HM, Nguyen CV, Nguyen TV, Nguyen MT, Thai HQ, Ho
489 MH, Thwaites G, Ngo HT, Baker S, Carrique-Mas J: **The use of colistin and**
490 **other critical antimicrobials on pig and chicken farms in southern Vietnam**
491 **and their association with resistance in commensal Escherichia coli. *Appl***
492 ***Environ Microbiol* 2016.**
- 493 12. Mulvey MR, Mataseje LF, Robertson J, Nash JH, Boerlin P, Toye B, Irwin R,
494 Melano RG: **Dissemination of the mcr-1 colistin resistance gene. *Lancet Infect***
495 ***Dis* 2016, 16:289-290.**
- 496 13. Grami R, Mansour W, Mehri W, Bouallegue O, Boujaafar N, Madec JY, Haenni
497 M: **Impact of food animal trade on the spread of mcr-1-mediated colistin**
498 **resistance, Tunisia, July 2015. *Euro Surveill* 2016, 21.**
- 499 14. Falgenhauer L, Waezsada SE, Yao Y, Imirzalioglu C, Kasbohrer A, Roesler U,
500 Michael GB, Schwarz S, Werner G, Kreienbrock L, et al: **Colistin resistance**

- 501 **gene mcr-1 in extended-spectrum beta-lactamase-producing and**
- 502 **carbapenemase-producing Gram-negative bacteria in Germany. *Lancet Infect***
- 503 *Dis* 2016, **16**:282-283.
- 504 15. Shen Z, Wang Y, Shen Y, Shen J, Wu C: **Early emergence of mcr-1 in**
- 505 ***Escherichia coli* from food-producing animals. *Lancet Infect Dis* 2016, **16**:293.**
- 506 16. Zhang R, Huang Y, Chan EW, Zhou H, Chen S: **Dissemination of the mcr-1**
- 507 **colistin resistance gene. *Lancet Infect Dis* 2016, **16**:291-292.**
- 508 17. Arcilla MS, van Hattem JM, Matamoros S, Melles DC, Penders J, de Jong MD,
- 509 Schultsz C, consortium C: **Dissemination of the mcr-1 colistin resistance gene.**
- 510 *Lancet Infect Dis* 2016, **16**:147-149.
- 511 18. Skov RL, Monnet DL: **Plasmid-mediated colistin resistance (mcr-1 gene):**
- 512 **three months later, the story unfolds. *Euro Surveill* 2016, **21**.**
- 513 19. Haenni M, Poirel L, Kieffer N, Chatre P, Saras E, Metayer V, Dumoulin R,
- 514 Nordmann P, Madec JY: **Co-occurrence of extended spectrum beta lactamase**
- 515 **and MCR-1 encoding genes on plasmids. *Lancet Infect Dis* 2016, **16**:281-282.**
- 516 20. Zhi C, Lv L, Yu LF, Doi Y, Liu JH: **Dissemination of the mcr-1 colistin**
- 517 **resistance gene. *Lancet Infect Dis* 2016, **16**:292-293.**
- 518 21. Zeng KJ, Doi Y, Patil S, Huang X, Tian GB: **Emergence of the Plasmid-**
- 519 **Mediated mcr-1 Gene in Colistin-Resistant *Enterobacter aerogenes* and**
- 520 ***Enterobacter cloacae*. *Antimicrob Agents Chemother* 2016, **60**:3862-3863.**
- 521 22. Du H, Chen L, Tang YW, Kreiswirth BN: **Emergence of the mcr-1 colistin**
- 522 **resistance gene in carbapenem-resistant *Enterobacteriaceae*. *Lancet Infect Dis***
- 523 **2016, **16**:287-288.**

- 524 23. Li A, Yang Y, Miao M, Chavda KD, Mediavilla JR, Xie X, Feng P, Tang YW,
525 Kreiswirth BN, Chen L, Du H: **Complete sequences of mcr-1-harboring**
526 **plasmids from extended spectrum beta-lactamase (ESBL)- and**
527 **carbapenemase-producing Enterobacteriaceae (CPE).** *Antimicrob Agents*
528 *Chemother* 2016.
- 529 24. Malhotra-Kumar S, Xavier BB, Das AJ, Lammens C, Butaye P, Goossens H:
530 **Colistin resistance gene mcr-1 harboured on a multidrug resistant plasmid.**
531 *Lancet Infect Dis* 2016, **16**:283-284.
- 532 25. Wood DE, Salzberg SL: **Kraken: ultrafast metagenomic sequence**
533 **classification using exact alignments.** *Genome Biol* 2014, **15**:R46.
- 534 26. Stoesser N, Sheppard AE, Pankhurst L, De Maio N, Moore CE, Sebra R, Turner
535 P, Anson LW, Kasarskis A, Batty EM, et al: **Evolutionary History of the Global**
536 **Emergence of the Escherichia coli Epidemic Clone ST131.** *MBio* 2016, **7**.
- 537 27. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin
538 VM, Nikolenko SI, Pham S, Prjibelski AD, et al: **SPAdes: a new genome**
539 **assembly algorithm and its applications to single-cell sequencing.** *J Comput*
540 *Biol* 2012, **19**:455-477.
- 541 28. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C,
542 Johnson JR, Walker AS, Peto TE, Crook DW: **Predicting antimicrobial**
543 **susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using**
544 **whole genomic sequence data.** *J Antimicrob Chemother* 2013, **68**:2234-2244.
- 545 29. Seemann T: **Prokka: rapid prokaryotic genome annotation.** *Bioinformatics*
546 2014, **30**:2068-2069.

- 547 30. Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD: **REAPR: a**
548 **universal tool for genome assembly evaluation.** *Genome Biol* 2013, **14**:R47.
- 549 31. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ: **IQ-TREE: a fast and**
550 **effective stochastic algorithm for estimating maximum-likelihood**
551 **phylogenies.** *Mol Biol Evol* 2015, **32**:268-274.
- 552 32. Didelot X, Wilson DJ: **ClonalFrameML: efficient inference of recombination**
553 **in whole bacterial genomes.** *PLoS Comput Biol* 2015, **11**:e1004041.
- 554 33. Wick RR, Schultz MB, Zobel J, Holt KE: **Bandage: interactive visualization of**
555 **de novo genome assemblies.** *Bioinformatics* 2015, **31**:3350-3352.
- 556 34. Woodford N, Turton JF, Livermore DM: **Multiresistant Gram-negative**
557 **bacteria: the role of high-risk clones in the dissemination of antibiotic**
558 **resistance.** *FEMS Microbiol Rev* 2011, **35**:736-755.
- 559 35. Snestrud E, He S, Chandler M, Dekker JP, Hickman AB, McGann P, Dyda F: **A**
560 **Model for Transposition of the Colistin Resistance Gene mcr-1 by ISAp11.**
561 *Antimicrob Agents Chemother* 2016.
- 562 36. Kusumoto M, Ooka T, Nishiya Y, Ogura Y, Saito T, Sekine Y, Iwata T, Akiba M,
563 Hayashi T: **Insertion sequence-excision enhancer removes transposable**
564 **elements from bacterial genomes and induces various genomic deletions.** *Nat*
565 *Commun* 2011, **2**:152.
- 566 37. Landersdorfer CB, Nation RL: **Colistin: how should it be dosed for the**
567 **critically ill?** *Semin Respir Crit Care Med* 2015, **36**:126-135.
- 568 38. Humphries RM: **Susceptibility testing of the polymyxins: where are we now?**
569 *Pharmacotherapy* 2015, **35**:22-27.

- 570 39. Carattoli A: **Resistance plasmid families in Enterobacteriaceae.** *Antimicrob*
571 *Agents Chemother* 2009, **53**:2227-2238.
- 572 40. Wong MH, Liu L, Yan M, Chan EW, Chen S: **Dissemination of IncI2 Plasmids**
573 **That Harbor the blaCTX-M Element among Clinical Salmonella Isolates.**
574 *Antimicrob Agents Chemother* 2015, **59**:5026-5028.
- 575 41. Lv L, Partridge SR, He L, Zeng Z, He D, Ye J, Liu JH: **Genetic characterization**
576 **of IncI2 plasmids carrying blaCTX-M-55 spreading in both pets and food**
577 **animals in China.** *Antimicrob Agents Chemother* 2013, **57**:2824-2827.
- 578 42. Liu L, He D, Lv L, Liu W, Chen X, Zeng Z, Partridge SR, Liu JH: **blaCTX-M-**
579 **1/9/1 Hybrid Genes May Have Been Generated from blaCTX-M-15 on an**
580 **IncI2 Plasmid.** *Antimicrob Agents Chemother* 2015, **59**:4464-4470.
- 581 43. Sun J, Li XP, Yang RS, Fang LX, Huo W, Li SM, Jiang P, Liao XP, Liu YH:
582 **Complete Nucleotide Sequence of IncI2 Plasmid Co-harboring blaCTX-M-55**
583 **and mcr-1.** *Antimicrob Agents Chemother* 2016.
- 584 44. Garcia-Fernandez A, Carattoli A: **Plasmid double locus sequence typing for**
585 **IncHI2 plasmids, a subtyping scheme for the characterization of IncHI2**
586 **plasmids carrying extended-spectrum beta-lactamase and quinolone**
587 **resistance genes.** *J Antimicrob Chemother* 2010, **65**:1155-1161.
- 588 45. Fang L, Li X, Li L, Li S, Liao X, Sun J, Liu Y: **Co-spread of metal and**
589 **antibiotic resistance within ST3-IncHI2 plasmids from E. coli isolates of**
590 **food-producing animals.** *Sci Rep* 2016, **6**:25312.
- 591 46. Yu H, Qu F, Shan B, Huang B, Jia W, Chen C, Li A, Miao M, Zhang X, Bao C, et
592 al: **Detection of mcr-1 colistin resistance gene in carbapenem-resistant**

593 **Enterobacteriaceae (CRE) from different hospitals in China. *Antimicrob***
594 *Agents Chemother* 2016.

595 47. Falgenhauer L, Waezsada SE, Gwozdziński K, Ghosh H, Doijad S, Bunk B,
596 Sproer C, Imirzalioglu C, Seifert H, Irrgang A, et al: **Chromosomal Locations of**
597 **mcr-1 and blaCTX-M-15 in Fluoroquinolone-Resistant Escherichia coli**
598 **ST410. *Emerg Infect Dis* 2016, 22.**

599 48. Tagg KA, Iredell JR, Partridge SR: **Complete sequencing of IncI1 sequence**
600 **type 2 plasmid pJIE512b indicates mobilization of blaCMY-2 from an**
601 **IncA/C plasmid. *Antimicrob Agents Chemother* 2014, 58:4949-4952.**

602 49. Tegetmeyer HE, Jones SC, Langford PR, Baltes N: **ISAp11, a novel insertion**
603 **element of Actinobacillus pleuropneumoniae, prevents ApxIV-based**
604 **serological detection of serotype 7 strain AP76. *Vet Microbiol* 2008, 128:342-**
605 **353.**

606 50. Stoesser N, Sheppard AE, Moore CE, Golubchik T, Parry CM, Nget P, Saroeun
607 M, Day NP, Giess A, Johnson JR, et al: **Extensive Within-Host Diversity in**
608 **Fecally Carried Extended-Spectrum-Beta-Lactamase-Producing Escherichia**
609 **coli Isolates: Implications for Transmission Analyses. *J Clin Microbiol* 2015,**
610 **53:2122-2131.**

611

612 **TABLES**

613 **Table 1. Resistance profiles of *mcr-I*-positive/cefotaxime-resistant study isolates.**

		Number of isolates resistant (%)	
All isolates (n=187)	Ampicillin	183	
	Amoxicillin-clavulanate	176	
Sequenced isolates (n=50)	Cefotaxime	165	
	Ceftazidime	131	
	Cefepime	139	
	Gentamicin	130	
	Amikacin	47	
	Ertapenem	23	
	Imipenem	4	
	Meropenem	16	
	Fosfomycin	101	
	Nitrofurantoin	19	
	Ciprofloxacin	152	
	Tigecycline	24	

615 **FIGURE LEGENDS**

616 **Figure 1.** Monthly proportions of *mcr-I*-positive human faecal samples in Guangzhou,
617 China, 2011-2016 for: Panel A, faecal samples harbouring *mcr-I*-positive isolates, and
618 Panel B, faecal samples harbouring *mcr-I*-positive/cefotaxime-resistant isolates. Black
619 line represents estimated prevalence by iterative sequential regression (ISR), with gaps
620 representing months with missing data. Blue lines at the base of the graph represent 95%
621 confidence intervals around the breakpoints estimated by the ISR model.

622
623 **Figure 2.** Phylogeny of sequenced *Escherichia coli* study isolates, plus available
624 reference sequences from NCBI (n=11), and associated sequence types (ST; “NF”
625 denotes “not found”); plasmid incompatibility groups; insertion sequences within 5kb of
626 *mcr-I* on either the same contig or contigs associated through the assembly graph; and
627 antimicrobial resistance genes present within the isolates (presence represented by
628 respective coloured shapes).

629
630 **Figure 3.** Alignment of *mcr-I* IncI2 study plasmid contigs (incomplete, denoted by *)
631 and circularized plasmid sequences, and three reference plasmid sequences (sequence
632 labels in bold followed by dates of host strain isolation). Pink/blue cross-links between
633 aligned sequences demonstrate regions of sequence homology (BLASTn matches of
634 ≥ 500 bp length, $>95\%$ sequence identity); dark regions within these cross-links
635 demonstrate sequence variation. Tick marks represent 10kb of sequence. Sequences are
636 ordered based on relatedness inferred from multiple sequence alignment (Kalign-EBI).

Coloured arrows demonstrate ORFs of particular relevance. Two plasmids were obtained from isolates cultured from the same patient (sequence labels in blue).

Figure 4. Alignment of *mcr-1* IncHI2/HI2A plasmid contigs and two reference plasmid sequences (sequence labels in bold followed by dates of host strain isolation). Pink/blue cross-links between aligned sequences demonstrate regions of sequence homology (BLASTn matches of ≥ 500 bp length, $>95\%$ sequence identity); dark regions within these cross-links demonstrate sequence variation. Tick marks represent 10kb of sequence. Coloured arrows demonstrate ORFs of particular relevance.

Figure 5. Alignment of *mcr-1* IncX4 study plasmid contigs (incomplete, denoted by *) and circularized plasmid sequences, and two reference plasmid sequences (sequence labels in bold followed by dates of host strain isolation; the top three sequences are incomplete plasmid contigs). Pink/blue cross-links between aligned sequences demonstrate regions of sequence homology (BLASTn matches of ≥ 500 bp length, $>95\%$ sequence identity); dark regions within these cross-links demonstrate sequence variation. Tick marks represent 10kb of sequence. Coloured arrows demonstrate ORFs of particular relevance.

Supplementary tables/figures

Table S1. Summary details of study isolate source, phenotype, lab results and whole-genome sequencing metrics.

660 **Figure S1.** Monthly counts of: (i) all faecal samples collected, (ii) faecal samples
661 harbouring *mcr-I*-positive isolates, and (iii) faecal samples harbouring *mcr-I*-
662 positive/cefotaxime-resistant isolates.

663

664 **Figure S2.** Alignment of the novel, untypeable *mcr-I* plasmid pSYSU0220_1 and incP
665 plasmid pHNFP671 (Accession number: KP324830).