Antimicrobial resistance and population genomics of extensively drug resistant *Escherichia coli* in pig farms in mainland China: a national wide epidemiological and microbiological study

Zhong Peng
Huazhong Agricultural University
Zizhe Hu
Huazhong Agricultural University
Zugang Li
Huazhong Agricultural University
Xiaoxue Zhang
Huazhong Agricultural University
Chaoying Jia
Huazhong Agricultural University
Tianzhi Li
Huazhong Agricultural University
Menghong Dai
Huazhong Agricultural University
Chen Tan
Huazhong Agricultural University
Zhuofei Xu
Shanghai MasScience Biotechnnology Institute
Bin Wu
Huazhong Agricultural University
Huanchun Chen
Huazhong Agricultural University
Xiangru Wang (wangxr228@mail.hzau.edu.cn)
Huazhong Agricultural University

Article

Keywords: Escherichia coli, pig farms. antimicrobial resistance phenotypes, population genomics; China

Posted Date: May 14th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-502741/v1

Version of Record: A version of this preprint was published at Nature Communications on March 2nd, 2022. See the published version at https://doi.org/10.1038/s41467-022-28750-6.

Abstract

Antimicrobial resistance (AMR) is one of the most urgent threats to the global public health, and the expanding use of antimicrobials in food animals is considered as a main reason for the worldwide rapid increasing of AMR. However, AMR in animals in many regions are poorly documented. China is the largest pig-rearing and pork consumption country in the world. In the present study, we identified AMR in pig farms from all provinces (including Tibet and Qinghai) of mainland China by investigation of a common indicator bacterium *Escherichia coli* from both pigs and the breeding environmental samples. A total of 2693 samples from pigs and environments in 67 pig farms in all 31 provinces of mainland China were collected between 1 October 2018 to 30 September 2019, and a total of 1871 *E. coli* strains were isolated. By testing the susceptibility of these 1871 *E. coli* isolates on 28 types of antibiotics that commonly used in both human and veterinary medicine, we found that resistance to tetracycline (96.26%), chloramphenicol (82.04%), moxifloxacin (81.56%), and trimethoprim/sulfamethoxazole (80.38%) were the broad phenotypes among these *E. coli* isolates from pig farms in China. A proportion of *E. coli* isolates were resistant to colistin (3.79%), carbapenems (imipenem [2.62%], meropenem [2.30%], ertapenem [2.46%]), and broad-spectrum-cephalosporins (ceftriaxone [29.56%], cefepime [14.00%]). More than 70% of the isolates displayed multidrug-resistant (MDR), and/or extensively drug-resistant (XDR) phenotypes, and MDR/XDR-*E. coli* was observed in pig farms in all provinces of mainland China. We also systematically revealed the distribution of O-serogroups, sequence types, resistance genes, virulence factors encoding genes, and putative plasmids of MDR/XDR-*E. coli* in pig farms from different provinces of China, and partially characterized the pathotypes of certain MDR/XDR-*E. coli* strains. In addition, the genetic transmission basis of the bla_{NDM}, *mcr*, ESBL-encoding, fluoroquinolone-resistance, and *tetX* genes were addressed in this study. Most importantly, we suggested a very high genetic propensity of the pig farm-sourced MDR/XDR-*E. coli* in spreading into humans. To the best of our knowledge, this is the first study on a national scale that the resistance phenotypes and population genomics of *E. coli* in pig farms in China are revealed. Our data presented herein will help understand the current profile of AMR in pigs and also provide reference for policy formulation of AMR control action in livestock in China.

Introduction

Antimicrobial resistance (AMR) is one of the most urgent threats to global public health. Recently, the emergence and rapid increase of multidrug-resistant (MDR), extensively drug-resistant (XDR), and even pan-drug-resistant (PDR) bacteria, particularly those resist the last-resort drugs (carbapenems, colistin, and tigecycline) have catalyzed the serious concern that the world may get its first inkling of impending catastrophe of a return to the pre-antibiotic [1–4]. Several drug-resistant bacteria have received a worldwide great concern, of particular note is *Escherichia coli* [1,5,6]. This bacterial species is not only a leading cause of foodborne infections, but also represents a major reservoir of antimicrobial resistance genes (ARGs) due to its great capacity to accumulate ARGs, mostly through horizontal gene transfer [7]. A recent study showed the total economic cost of AMR in *Staphylococcus aureus*, *E. coli*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, and *Pseudomonas aeruginosa* reached $2.8 billion per year in the United

States [8] . In addition, *E. coli* has been found to play important roles in disseminations of *bla* $_{NDM-1}$, *mcr*, and/or *tet*(X3)/(X4); these ARGs mediates resistance to the last-resort drugs (carbapenems, colistin, and tigecycline) for gram-negative bacteria which may lead to non-antimicrobials available in both human and veterinary medicine [3,4,9] . In addition, extended-spectrum beta-lactamase (ESBL)-producing *E. coli* is also a major global health problem [1,5] .

The large and expanding use of antimicrobials in food animals is considered as a main reason for the worldwide rapid increasing of AMR [10]. In the past two decades, meat production has plateaued in high-income countries but has largely grown in low- and middle-income countries particularly in Asia [11]. Under this background, antimicrobials are widely used in livestock either for growth promotion and/or for health maintain. Globally, 73% of all antimicrobials sold on the earth are used in animals raised for food [11]. The large and expanding use of antimicrobials can lead to the development of drug-resistant bacteria in the animal guts, and these resistant bacteria can spread to humans [1]. To date, AMR in animals in low- and middle-income countries are poorly documented, and this may be in part due to the absence of systematic surveillance systems and large epidemiological studies [11].

China is the largest developing country in the world. Many provinces in China still belong to the low- and middle-income regions. China is also the number one country for antimicrobials producing, pig rearing and production, and pork consumption in the world [12,13]. Therefore, understanding the current profile of AMR in pig farms in China has a great significance for global AMR surveillance. However, pig producing in China is very complex with different sizes of pig farms and production models in various regions. To date, there is still lack of systemic data reflecting the condition of AMR in the pork production chain in China. Here, we identified AMR in pig farms in most regions in China by investigation of *E. coli*, which is a commonly used biomarker of AMR in pig farms [14]. To the best of our knowledge, this is for the first time that the pig farms in all regions in mainland China were included for the epidemiological investigation. This study will contribute to understand the current profile of AMR in pigs and also provide reference for policy formulation of AMR control action in livestock in China.

Results

Antimicrobial resistance phenotypes and microbiological characteristics of *E. coli* isolates from pig farms in China. Between 1 October 2018 to 30 September 2019, a total of 2693 samples from pigs and breeding environments in 67 pig farms in all 31 provinces of mainland China were collected for *E. coli* isolation, and a total of 1871 *E. coli* strains including 1108 strains from pig samples (anal swabs and/or diarrheal feces) and 763 strains from environmental samples (swabs of drinking and/or fecal slurry, floors, drinking and/or food troughs) were finally obtained (Fig. 1A). Initially, we attempted to include more farms in each of the provinces, but the sudden outbreak of African Swine Fever (ASF) in late 2018 and the coronavirus disease (COVID-19) in late 2019 made it extremely difficult to get more samples from more farms. By testing the minimum inhibitory concentrations (MICs) of 28 types of antimicrobials belonging to aminoglycosides (amikacin [AMK], gentamicin [GEN], tobramycin [TOB]), carbapenems

(imipenem [IPM], meropenem [MRP], ertapenem [ETP]), cephalosporins (cefazolin [CFZ], cefuroxime [CFX], cefoxitin [FOX], ceftazidime [CAZ], ceftriaxone [CRO], cefepime [CPM]), β-lactam combination agents (amoxicillin/clavulanate [AMC], ampicillin/sulbactam [AMS], piperacillin/tazobactam [PTZ]), monobactams (aztreonam [AZM]), phenicols (chloramphenicol [CHL]), tetracyclines (tetracycline [TET], minocycline [MIN], tigecycline [TGC]), fluoroquinolones (moxifloxacin [MXF], ciprofloxacin [CIP], levofloxacin [LVX], norfloxacin [NOR]), sulfonamides (trimethoprim/sulfamethoxazole [SXT]), fosfomycins (fosfomycin [FOS]), nitrofurantoins (nitrofurantoin [NIT]), and polymyxins (colistin [CL]) on all these 1871 isolates, the resistant phenotypes of *E. coli* strains in China were characterized (Figs. 1B **& C**). The results revealed that resistance to TET (percent resistant strains: 96.26%, *n* = 1801), CHL (82.04%, *n* = 1535), MXF (81.56%, *n* = 1526), and SXT (80.38%, *n* = 1504) were broad phenotypes of *E. coli* isolates from pig farms in China (Fig. 1C). In particularly, a number of *E. coli* isolates from pig farms were resistant to colistin (3.79%, *n* = 71), carbapenems (IPM [2.62%, *n* = 49], MRP [2.30%, *n* = 43], ETP [2.46%, *n* = 46]), and broad-spectrum-cephalosporins (CRO [29.56%, *n* = 553], CPM [14.00%, *n* = 262]) (Fig. 1C). Many isolates showed resistance to TGC (37.31%, *n* = 698), but most of them with MIC values ranging from 0.5 μg/ml to 1 μg/ml (92.98%, *n* = 649), only a very small proportion of them showed high-level of resistance (MIC ≥ 4 μg/ml; 0.72%, *n* = 5) (Figs. 1B **& C**). When comparing isolates from different samples of collection, resistance rates of isolates from pigs against SXT, GEN, CIP and NOR were significantly higher than those of the isolates from environmental samples, while isolates from environmental samples had significantly higher resistance rates against TGC, AMS, and AMC compared to isolates from pigs (Supplementary materials Figure S1).

The *E. coli* isolates from pig farms in different regions in China showed serious phenotypes of multidrug resistance and even extensively drug-resistance. More than 70% of the isolates were resistant to more than three of the twelve (sub-)classes (aminoglycosides, carbapenems, cephalosporins, β-lactam combination agents, monobactams, phenicols, tetracyclines, fluoroquinolones, sulfonamides, fosfomycins, nitrofurantoins, and polymyxins) of antibiotics tested (Figs. 2A **& B**). The mainland China consists of a total of 31 provinces, and MDR-*E. coli* isolates were determined in all 31 provinces, while XDR-*E. coli* isolates were determined in 29 provinces (except Sichuan and Tibet; Figs. 2C **~ E**). In Sichuan and Tibet, 69.44% and 42.86% of the isolates from pig farms were resistant to more than three (sub-)classes of antibiotics, respectively. Notably, in Beijing and Ningxia, all isolates from pig farms displayed multidrug resistance (Figs. 2D **& E**). In particular, carbapenem-resistance isolates were found in 7 provinces but most of them were isolated from Henan province; isolates resistant to colistin were found in 12 provinces and most of them were also recovered from Henan province (Figs. 2B **& C**). TGC-resistant *E. coli* were found in 28 provinces including Tibet (Fig. 2D), but high-level TGC-resistant strains (MIC value ≥ 4 μg/ml) were only determined in Anhui, Hunan, Guizhou, Hebei and Hubei (Fig. 2C). *E. coli* strains resistant to broad-spectrum-cephalosporins (CRO and CPM) were isolated from pig farms in 30 and 24 provinces, respectively (Figs. 2C **& D**). Tibet was the only region where no strains from pig farms with the above particularly mentioned resistant phenotypes being detected (Figs. 2C **& D**).

Determination of putative pathogenic XDR-*E. coli* isolates. To understand the genomic characteristics of drug-resistant *E. coli* isolates from pig farms in different regions of China, we selected isolates (totally

515) with resistant phenotypes to either carbapenems (*n* = 49), CL (*n* = 71), TGC (MIC value ≥ 4 μg/ml; *n* = 5), or broad-spectrum-cephalosporins (*n* = 495) for next-generation sequencing (NGS). These isolates also displayed resistance to aminoglycosides (*n* = 334), phenicols (*n* = 476), tetracyclines (*n* = 510), fluoroquinolones (*n* = 473), sulfonamides (*n* = 452), and/or nitrofurantoins (*n* = 59) (Fig. 3A; Supplementary materials Table S1). *In silico* serotyping using the whole genome sequences identified 101 kinds of O-serogroups for the sequenced isolates, and O9a (*n* = 50), O101 (*n* = 46), and O8 (*n* = 39) were the predominant types (Fig. 3A; Supplementary materials Table S1). Notably, many O-serogroups which might have public health significance were also determined, including O101, O128ac, O11, O136, O28ac, O103, O149, O15, O45, O125ab, O9, O115, O159, O73, O25, O26, O29, O6, O8, O80, O143, O148, O153, O157, O166, O167, O78, O86, and O91 (Fig. 3A; Supplementary materials Table S1). Multilocus sequence typing (MLST) analyses revealed that these MDR-isolates belonged to 118 different sequence types (STs) and ST10 (*n* = 52), ST101 (*n* = 39), ST48 (*n* = 22) as well as ST5229 (*n* = 20) were the broadly determined STs (Fig. 3B). Interestingly, there was little or no correlation between the phylogenetic groups, STs, O-serotypes, and the place of isolation (Supplementary materials Table S1).

We next analyzed the virulence factors encoding genes (VFGs) carried by the MDR/XDR-*E. coli* isolates in this study. According to prediction, each of the isolates contained numerous VFGs (numbers of VFGs ranged from 78 to 284; Supplementary materials Table S2). Of particularly note were *astA*, *eae*, *east1*, *ecpABCDER*, *efa1*, *eltAB*, *escCDFJNRSTUV*, *espABD*, *estIa*, *paa*, *pic*, *stb*, stx_{2eB}, *tir*, *toxB* (Fig. 3A; Supplementary materials Table S2). These genes encode important adherence factors and/or toxins, and *E. coli* isolates possessing these VFGs are assigned as pathogenic *E. coli* [15]. In particular, these VFGs were determined in MDR/XDR-*E. coli* isolates with O-serotypes (O6, O8, O9, O11, O15, O25, O26, O28ac/O42, O29, O45, O73, O78, O80, O86, O91, O101, O103, O115, O125ab, O128ac, O136, O143, O148, O149, O153, O157, O159, O166, O167) that have public health significance (Fig. 3A).

Genomic associations with antimicrobial resistance phenotypes. We first searched for the presence of known acquired antimicrobial resistance genes (ARGs) in the genomic sequences. This approach led to the detection of 109 kinds of ARGs, including 28 aminoglycoside resistance genes, 35 β-lactam resistance genes, 7 phenicol resistance genes, five tetracycline resistance genes, ten quinolone resistance genes, as well as three sulfonamide resistance genes, seven macrolide resistance genes, two rifampicin resistance genes, seven trimethoprim resistance genes, two fosfomycin resistance genes, and four lincosamide resistance genes (Supplementary materials Table S3). In particular, $bla_{CTX-M-55}$ and bla_{TEM-1B} were most-frequently determined ESBL genes in *E. coli* isolates ($bla_{CTX-M-55}$ was carried in 209 sequenced isolates while bla_{TEM-1B} was carried in 181 ones) from pig farms in China (Fig. 4A); while *qnrS1*, *oqxB*, and *oqxA* were most-frequently determined quinolone resistance genes (presence in 233, 114, and 109 sequenced isolates, respectively; Fig. 4B). Among the tetracycline-resistant isolates, *tet(A)* and *tet(M)* were most-frequently determined resistance genes (presence in 453 and 170 sequenced isolates, respectively; Fig. 4C). Notably, three carbapenem resistance genes (bla_{NDM-1}, bla_{NDM-5}, bla_{NDM-7}), two colistin resistance genes (*mcr-1.1*, *mcr-3.1*), and one tetracycline resistance genes (*tetX4*) were determined, and they conferred resistance to carbapenems, colistin, and tigecycline, respectively

(Figs. 4C ~ **E**). Over 15% NDM-producing isolates carried colistin resistance gene *mcr-1*, while only one isolates carried both colistin resistance gene *mcr-1* and high-level tetracycline resistance gene *tetX4*. We also observed there was a strong correlation between the presence of ARGs and the expected resistance phenotypes. Only carbapenem resistance isolates harboring the three carbapenem resistance genes, while only isolates with high level resistance to tigecycline (MIC value ≥ 4 μg/ml) containing *tetX4*; in addition, the *mcr* genes were only found in isolates with resistance phenotypes to colistin (Fig. 4).

We also detected the point mutations associated with AMR in the *E. coli* isolates from pig farms. Point mutations were determined in *gyrA*, *gyrB*, *parC*, *parE*, *pmrA*, *pmrB*, *folP*, *ampC*, *rpoB*, 23SrRNA, 16S_rrsB, 16S_rrsC, and 16S_rrsH (Supplementary materials Table S4). In addition to RNA mutations observed in 23SrRNA, 16S_rrsB, 16S_rrsC, and 16S_rrsH, point mutations in *parC* were determined in most of the pig farm *E. coli* isolates in China (*n* = 514) while mutations in *ropB* were determined in small numbers of *E. coli* isolates (*n* = 22) (Fig. 5A). Point mutations caused 31 types of amino acid changes in GyrA, and "S83L" (*n* = 301) as well as "D87N" (*n* = 203) were the most common mutations (Fig. 5B). In GyrB, point mutations caused 21 types of amino acid changes, and "A618T" (*n* = 18), "E219K" (*n* = 18), "H652R" (*n* = 17), and "S492N" (*n* = 15) were the most common mutations (Fig. 5C). Point mutations caused 39 types of amino acid changes in ParC, and "E62K" (*n* = 514) as well as "S80I" (*n* = 237) were the most common mutations (Fig. 5D). In ParE, point mutations caused 29 types of amino acid changes, and "S458A" (*n* = 53) was the most common mutation (Fig. 5E). In PmrA, point mutations caused 11 types of amino acid changes, and "G144S" (*n* = 30) was the most common mutation (Fig. 5F). Point mutations caused 22 types of amino acid changes in PmrB, and "D283G" (*n* = 247), "Y358N" (*n* = 224) as well as "H2R" (*n* = 109) were the most common mutations (Fig. 5G). In FolP, point mutations caused 14 types of amino acid changes, and "F4Y" (*n* = 20) was the common change (Fig. 5H). In AmpC, point mutations caused 23 types of amino acid changes, and deletion of arginine at site 24 (R24*; *n* = 18) was the common change (Fig. 5I). In RpoB, point mutations caused 19 types of amino acid changes (Fig. 5J)

Genetic basis of the ESBL, fluoroquinolone-resistance, bla_{NDM}, mcr, and tetX genes transmission. To understand the genetic basis of the ARG transmission, we next analyzed the presence of ARG-associated plasmids. This approach led to the detection of 53 groups of plasmid replicons (Supplementary materials Table S5). We found a ColRNAI type plasmid was presented in most of the *E. coli* isolates (*n* = 272) from pig farms in China, followed by IncFIB (*n* = 266), and IncX1 (*n* = 236) (Fig. 6A). We determined a total of 53 groups of plasmids in ESBL-genes carrying isolates (*n* = 495), quinolone-resistance-genes carrying isolates (*n* = 473), and all *tet*-carrying isolates (*n* = 510) from pig farms in China, respectively; and each of the isolates contained 0 to 14 of the reference plasmids (Figs. 4A ~ **C;** Supplementary materials Table S5). Plasmids including ColRNAI, IncFIB, and IncX1 were observed in most of the ESBL-genes carrying isolates [ColRNAI (260/495), IncFIB (257/495), IncX1 (228/495)], quinolone-resistance-genes carrying isolates [ColRNAI (248/473), IncFIB (257/473), IncX1 (228/473)], as well as the *tet*-carrying isolates [ColRNAI (270/510), IncFIB (265/510), IncX1 (235/510)] from pig farms in China. However, plasmids including Col(MP18), Col3M, IncB/O/K/Z, and IncU were rarely determined in these isolates. Short-read mapping of all bla_{NDM}-carrying isolates from pig farms in China (*n* = 45) determined 34 groups of

plasmids, and each of the isolates carried 2 to 14 of the reference plasmids (Fig. 4D; Supplementary materials Table S5). Most *bla*$_{NDM}$-carrying isolates carried IncFII(pHN7A8) (27/45), ColRNAI (25/45), IncX3 (24/45), Col(MG828) (21/45), and Col156 (20/45). However, some plasmids such as Col(BS512), Col3M, IncFIB(Mar), IncFIB(pHCM2), IncFII(pCoo), and IncP1 were found only rarely. A total of 41 groups of plasmids were determined in the *mcr*-carrying isolates (n = 69), and each of the isolates carried 3 to 14 of the reference plasmids (Fig. 4E; Supplementary materials Table S5). Most *mcr*-carrying isolates carried ColRNAI (50/69), IncX1 (39/69), RepA (36/69), IncHI2 (36/69), and IncHI2A (36/69); while several ones including Col(KPHS6), Col(Ye4449), IncFII(pCoo), and IncP1 were rarely found.

We generated the complete genome sequences of the *bla*$_{NDM}$-, *mcr*-, and/or *tetX4*-carrying plasmids by Oxford Nanopore Sequencing (ONT). We obtained an 85.9-kb IncFII-type-*bla*$_{NDM-1}$-carrying plasmid (designated pXD33-05) and a 33.3-kb IncX4-type-*mcr-1*-carrying plasmid (designated pXD33-06) from an isolate XD33 co-producing NDM and MCR (GenBank accession no. JAENDM000000000). Interestingly, sequence alignments revealed that a pXD33-05-like plasmid and a pXD33-06-like plasmid were also presented in the other isolates co-producing NDM and MCR (Supplementary materials Figure S2). Sequence alignment also revealed that *bla*$_{NDM-1}$-carrying plasmid pXD33-05 was highly homologous to a *bla*$_{NDM-1}$-carrying plasmid pHNEC55 (GenBank accession no. KT879914) (Fig. 6B). The average nucleotide identity (ANI) between the backbones of pXD33-05 and pHNEC55 was higher than 99%. However, the MDR elements between the two plasmids were different. Structurally, the MDR elements of pXD33-05 consisted of two ARG cassettes, including a 7.6-kb cassette harboring a bleomycin resistance gene, a aminoglycoside resistance gene *aph(3')-VI* and *bla*$_{NDM-1}$ as well as a 2.4-kb one harboring a aminoglycoside resistance gene *rmtB* and a ESBL-encoding gene *bla*$_{TEM-1B}$ (Fig. 6B). The 7.6-kb cassette was flanked by an *IS6* and an *IS3* elements, while the 2.4-kb cassette was flanked by a Tn3 and an *IS6* elements. Sequence comparisons showed that the *mcr-1*-carrying plasmid pXD33-06 was highly homologous to plasmid pWI2-mcr (GenBank accession no. LT838201) (Fig. 6C). However, it displayed little homology to the high-impact *mcr*-bearing plasmid pHNSHP45 (GenBank accession no. KP347127) reported in China [4] (Fig. 6C). In addition to *mcr-1*, no other ARGs were found on pXD33-06 (Fig. 6C). We also analyzed the genetic environments of *tetX4* which mediates resistance to high level tigecycline in *E. coli* isolates from pig farms in China. ONT sequencing on two high-level-TGC resistant isolates HB50 and SY36 revealed that *tetX4* was carried by an IncX1 plasmid in both isolates, which was highly homologous to a previously reported *E. coli tetX4*-harboring plasmid pYY76-1-2 (GenBank accession no. CP040929) [16] (Fig. 6D). In all determined *tetX4*-carrying elements, *tetX4* was adjacent to an IS*CR2* element (Fig. 6D). Plasmid conjugation experiments revealed that most of the *bla*$_{NDM}$-, *mcr*-, and/or *tetX4*-carrying plasmids were conjugative and conferred phenotypes of carbapenem-, colistin-, and high-level tigecycline (MIC value ≥ 4 μg/ml) resistance to the bacterial receipts, respectively (Supplementary materials Table S6).

High genetic propensity of farm sourced XDR-*E. coli* in spreading into humans. To determine the genetic propensity of the MDR/XDR-*E. coli* isolates to spread into the human sector, the genetic relatedness of the 515 MDR/XDR-*E. coli* isolates from pig farms in China to 287 publicly available draft genomes of human commensal *E. coli* (Bioproject no. PRJNA4001047) were investigated [17]. The 802 *E. coli* isolates were

phylogenetically divided into three lineages (Fig. 7A), and the 515 MDR/XDR-*E. coli* isolates from pig farms displayed a close relatedness to the 287 human *E. coli* strains (Fig. 7B). A large proportion of pig-farm originated MDR/XDR-*E. coli* isolates showed high genetic similarity (443/515, differed by only less than 1000 SNPs) to the human originated *E. coli* strains in China (Fig. 7; Supplementary materials Table S7). Of particularly concern is that 44.27% (228/515) of the MDR/XDR-*E. coli* isolates from pig farms differed by only less than 100 SNPs (as small as 3 SNPs) from the human *E. coli* isolated in China (Fig. 7; Supplementary materials Table S7).

Discussion

In this study, we investigated AMR phenotypes of *E. coli* isolated from both pig samples and environmental samples in pig farms from all provincial regions in mainland China. As an important natural reservoir of ARGs and a commonly-used biomarker bacteria for monitoring AMR [7,14], the resistance profile of *E. coli* from pig farms may reflect the AMR condition in these farms. To the best of our knowledge, this is the first time the AMR phenotypes of *E. coli* isolates from pig farms in all provinces of mainland China, including Tibet and Qinghai where there were little related data before [11] being reported.

Our determination of resistance phenotypes of *E. coli* isolates from pig farms in different provinces in China on 28 types of antimicrobials commonly used in both human and veterinary medicine indicated a worrisome condition of AMR in pig farms (Figs. 1&2). MDR and even XDR were the common phenotypes of *E. coli* isolates recovered in this study, and MDR/XDR-*E. coli* isolates were widely determined in pig farms in different provinces in whole mainland China, including Tibet, Xinjiang, and Qinghai (Fig. 2). This worrisome condition is widely accepted as the result of overuse and abuse of antibiotics in pig industry in the country [12,18]. During the past decades, along with a rapid increase in economy, the production of meat, eggs, and milk in China has rapidly increased, especially for pork, the main source of animal protein for most Chinese people [12,19]. To meet increasing demand for pork, both the number and the size of pig farms have grown markedly, and a massive amount of antibiotics are used in the country to support its rapid increase in pig production [12]. According to available reports [13,20,21], major classes of antibiotics extensively used in China's pig farms include sulphonamides, tetracyclines, fluoroquinolones, macrolides, and β-lactams, and in particular, fluoroquinolones and β-lactams contributed more than half. Consistently, resistance to antibiotics belonging to those antibiotic classes were found to be the broad phenotypes for *E. coli* isolates from pig farms in China (Fig. 2).

Among the drug-resistant *E. coli* determined in pig farms in China, of particular note are the isolates with resistance phenotypes to carbapenems, CL, and TGC (Fig. 2). In particular, over 15% NDM-producing isolates carried CL resistance gene *mcr-1*. All of these antibiotics are proposed as the last-resort antibiotics for treating infections caused by MDR Gram-negative bacteria [4,22]. According to the recent official policy, the carbapenem antibiotics are not approved to be used in livestock in China, and it remains to be elucidated why *E. coli* isolates with resistance phenotypes to these antibiotics are

recovered from pig farms. Notably, bacterial isolates with resistance phenotypes to carbapenems have also be recovered from poultry farms in China [23]. Although we cannot exclude the possibility that some pig farms might secretly use carbapenem antibiotics without receiving approval, a more preferrable possibility for the acquisition of these phenotypes is due to contaminated in-house environment [23]. The use of carbapenems in Chinese hospitals might cause the contamination of carbapenem-resistant bacteria or mobile carbapenem-resistant genes (e.g., bla_{NDM}) in environments (e.g., water, air, etc.). These resistant bacteria or ARGs may spread to livestock farms and in turn, lead to the emergence of carbapenem-resistant bacteria in animals. It was worth note that most of the carbapenem-resistant *E. coli* isolates recovered from pig farms in China in this study carried the NDM-producing gene bla_{NDM} (Fig. 4D). Continuous monitoring the persistence and spread of carbapenem-resistant bacteria in animals particularly food producing animals is essential, as these bacteria may represent high risks to human health.

The recovery of CL-resistant *E. coli* from pig farms might also due to the acquisition of CL-resistant *mcr* gene from the environment. Although the colistin withdrawal policy in 2017 and the decreasing use of colistin in agriculture have had a significant effect on reducing colistin resistance in both animals and humans in China [24], plasmid-mediated CL resistance *mcr* genes might also persist in the environments of livestock farms [23,25]. The persistence of the *mcr* genes may contribute to the dissemination of CL-resistant bacteria. In agreement with this hypothesis, the plasmid-mediated colistin resistance genes *mcr-1* and *mcr-3* were widely determined in *E. coli* isolates with resistance phenotypes to CL (Fig. 4E). Notably, CL-resistant *E. coli* were found in pig farms in 11 provinces of China in this study, suggesting a still worrisome condition for the persistence and dissemination of CL resistance. More active actions should be taken to solve the problem and continuous monitoring is still necessary [24]. Although a large number of *E. coli* isolates from pig farms in China showed resistance to TGC (37.31%, $n = 698$), most of them displayed low-level of resistance (MIC values ranging from 0.5 μg/ml to 1 μg/ml; 92.98%, $n = 649$), and only five isolates displayed high-level of resistance (MIC ≥ 4 μg/ml; 0.72%) (Figs. 1B **& C**). However, the previously reported high-level TGC conferring gene *tetX4* [3] was only found in high-level TGC resistant isolates. Instead, several *tet* genes were widely detected in low-level TGC resistant isolates (Figs. 1B **& C**). Phenotypes of low-level TGC resistance might be conferred by these *tet* genes such as *tetA*, *tetB*, *tetC*, and/or *tetM* [26].

By performing whole genome sequencing, the population genomics of MDR/XDR-*E. coli* in pig farms in China was revealed. During the analyses, we characterized several O-serogroups that are frequently associated with *E. coli* pathotypes including Enterotoxigenic *E. coli* (ETEC), Enteropathogenic *E. coli* (EPEC), Enterohaemorrhagic *E. coli* (EHEC), Enteroaggregative *E. coli* (EAEC), Enteroinvasive *E. coli* (EIEC), and Diffusely adherent *E. coli* (DAEC) [27,28] (Fig. 3). Importantly, several marker VFGs of *E. coli* pathotypes (e.g., *astA*, *eae*, *east1*, *stb*, stx_{2eB}, *toxB*, etc.[27,28]) were also determined in *E. coli* isolates belonging to these O-serogroups (Fig. 3). The wide presence of these MDR/XDR-*E. coli* in pig farms represents a great health risks and should receive more attentions. Determination of ARGs identified mobile genes conferring resistance to the tested antibiotic classes in *E. coli* isolates in pig farms in China (Fig. 3; Supplementary

materials Table S3). In addition, a large number of macrolide-resistance genes were also determined. These findings suggest that the extensive use of sulphonamides, tetracyclines, fluoroquinolones, macrolides, and β-lactams in China's pig farms facilitates the acquisition of related resistance genes in *E. coli* in farms. Along with the detection of multiple ARGs, we also detected many putative plasmids in *E. coli* isolates from pig farms in China (Supplementary materials Table S4). Several groups of these plasmids (e.g., IncA/C, IncB/O, IncFIB, IncX1, IncFIIA, IncX2, IncY, etc.) are capable of carrying transfer, MDR, and virulence functions in broad-host-range [29]. The presence of those plasmids may accelerate the dissemination of ARGs and even VFGs as well. We also determined several groups of plasmids that were rarely associated with the spread of specific ARGs. For example, the IncX3 plasmid has been reported to be account for majority of bla_{NDM} carriage in livestock farms [23,30]. However, we determined an IncFII-type-bla_{NDM}-carrying plasmid (Fig. 6B), and this type of plasmid was observed in all *E. coli* isolates co-producing NDM and MCR (Supplementary materials Figure S2). In addition, the *mcr*-carrying plasmid determined *E. coli* isolates co-producing NDM and MCR in this study was also different from the plasmid mediating the dissemination of *mcr* in China reported recently [4] (Fig. 6C). The presence of these plasmids makes the dissemination of bla_{NDM-1} and/or *mcr* more heterogeneous. It was worth note that the elements mediating the spread of high-level TGC resistance gene *tetX4* in *E. coli* isolates from pig farms in this study shared highly homologous to those reported previously [3,17,31] (Fig. 6D), suggesting the dissemination of *tetX4* might be not as heterogeneous as bla_{NDM} and *mcr*. However, continuous monitoring should be taken in the future. Most notably, the MDR/XDR-*E. coli* isolates from pig farms displayed a very close relatedness to the *E. coli* strains from humans in China (Fig. 7). Most pig farm-origin *E. coli* isolates in this study differed by only less than 1000 SNPs to the human-originated *E. coli* strains, and many differed by only less than 10 SNPs (Fig. 7; Supplementary materials Table S7). These findings suggest a very high genetic propensity of farm sourced MDR/XDR-*E. coli* in spreading into humans in China.

Conclusions

Although this work has a limitation that we could not include more pig farms in different provinces in China for sample collection due to the outbreak of ASF and COVID-19, our sample collection still covers pig farms in all provinces of China. On a national scale for the first time, we characterized the resistance phenotypes as well as population genomics of *E. coli* in pig farms in China. Our results revealed a worrisome condition of AMR in pig farms in China and there is still a long way for China to take actions to reduce AMR in livestock. Fortunately, Chinese government has taken a series of active actions to solve the urgent AMR conditions in animal husbandry. A noteworthy action is the Ministry of Agriculture and Rural Affairs (MARA) has issued a policy to ban the addition of antibiotics for promoting animal growth in feed from July 1, 2020 (MARA Announcement No.194, 07-10-2019). In this study, we also systematically revealed the distribution of O-serogroups, sequence types, ARGs, VFGs, as well as putative plasmids of MDR/XDR-*E. coli* in pig farms in different provinces of China. These data will provide comprehensive insights to help understand AMR in pig farms, and may also be beneficial for the government to make policies for reducing AMR in pig industry in the country. Notably, we also determined

many MDR/XDR-*E. coli* with potential pathogenicity to humans and most importantly, we found there was a very high genetic propensity of pig farm-sourced MDR/XDR-*E. coli* in spreading into humans. The persistence and dissemination of these isolates represent important health risks and should receive more attentions.

Materials And Methods

Sample collection, identification, and antimicrobial susceptibility testing

Between 1 October 2018 to 30 September 2019, a surveillance project on AMR in *E. coli* from pig farms in China was set. In each of the 31 provinces in mainland China, 2 ~ 3 different pig farms were randomly selected to collect swabs from fresh feces and rectal swabs of pigs (with approximately 40 samples per farm), as well as swabs of drinking and fecal slurry, floors, troughs (for each point at least three samples per farm were collected). All samples were shipped with dry ice immediately for *E. coli* isolation. Swabs were incubated in Luria Bertani (LB) broth (Sigma-Aldrich, MO, USA) at 37°C, 180 rpm, for overnight. Afterwards, swab cultures were streaked on MacConkey agars and were incubated at 37°C for 16 hours. *E. coli* isolates were confirmed by PCR amplification of the 16S rRNA gene with primers (F: 5'-GAAGCTTGCTTCTTTGCT-3', R: 5'-GAGCCCGGGGATTTCACAT-3') documented previously [32]. PCR assays amplifying the presence of seven house-keeping genes of *E. coli* (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) [33] were set for double confirmation.

The antimicrobial susceptibility of *E. coli* isolates was determined by testing the MIC values of different kinds of antibiotics on the bacterium, following the Clinical & Laboratory Standards Institute (CLSI, United States) recommended microbroth dilution protocol (CLSI M100, 28th Edition). A total of 28 types of antibiotics (AMK, GEN, TOB, IPM, MRP, ETP, CFZ, CFX, FOX, CAZ, CRO, CPM, AMC, AMS, PTZ, AZM, CHL, TET, MIN, TGC, MXF, CIP, LVX, NOR, SXT, FOS, NIT, CL) were included for the tests. Results were interpreted using the CLSI breakpoints (CLSI M100, 28th Edition). If CLSI breakpoint was not available, the European Committee on Antimicrobial Susceptibility Testing (EUCAST) breakpoint (v 8.1) was alternatively used for the interpretation. Each antibiotic was tested with three duplicates. *E. coli* ATCC 25922 was used as quality control.

Whole genome sequencing and data availability

All *E. coli* strains with resistant phenotypes to one of the carbapenems tested, broad-spectrum-cephalosporins (ceftriaxone and cefepime), tigecycline, or colistin, were selected for NGS. Bacterial genomic DNA was extracted by using a commercial DNA Kit (TIANGEN, Beijing, China). The quality and concentration of the bacterial genomic DNA was evaluated via electrophoresis on a 1% agarose gel as well as NanoDrop2000 (Thermo Scientific, Waltham, MA, USA) and Qubit 4 Fluorometer (Thermo Scientific, Waltham, USA). Libraries were constructed based on the qualified DNA by using a NEBNext Ultra™ II DNA Library Prep Kit (New England BioLabs, Ipswich, USA), and were sequenced on a NovaSeq 6000 platform using the pair-end 150-bp sequencing protocol (Novogene, Beijing, China). Raw reads with

low quality were removed as described previously [34]. High-quality reads were *de novo* assembled via SPAdesv3.9.0 to generate genome contigs.

The complete genome sequences of bla_{NDM-1}-, *mcr-1*-, and/or *tetX4*-carrying plasmids were generated by ONT sequencing in combination with the Illumina technology. Plasmid DNA was extracted using the phenol-chloroform protocol combined with Phase Lock Gel tubes (Qiagen GmbH) and was detected by the agarose gel electrophoresis as well as quantified by Qubit® 2.0 (Thermo Scientific, Waltham, USA). Libraries for ONT and Illumina sequencing were prepared using an SQK-LSK109 kit and a NEBNext® Ultra™ DNA Library Prep kit, respectively. Prepared DNA libraries were sequenced using Nanopore PromethION platform and Illumina NovaSeq PE150 at Novogene Co. LTD (Tianjin, China), respectively. ONT and Illumina short reads were finally assembled and combined using the Unicycler v0.4.4 software with default parameters.

Whole genome sequences (WGSs) of *E. coli* isolates have been deposited into GenBank (BioProject accession no. PRJNA688628). GenBank accession numbers are given in Supplementary materials Table S1.

Bioinformatical analysis

ResFinder 4.1 [35] was used to determine putative acquired antimicrobial resistance genes (ARGs) and point mutations associated with AMR. PlasmidFinder 2.1 [36] was used to determine putative plasmids carried by the sequenced strains. *In silico* serogroups and sequence types (ST) of the sequenced strains were determined by SerotypeFinder 2.0 [37] and multilocus Sequence Typing (MLST) 2.0 [38], respectively. Sequence alignments were performed by using the MAFFT software version 7.471 [39]. RAST Sever was used for sequence annotations [40]. Average nucleotide identities between two genome sequences were calculated by ANI calculator [41]. A comparative genome analysis was performed and visualized using the BRIG package [42] and/or the EasyFig package [43]. Phylogenetic trees based on whole-genome single nucleotide polymorphisms (WG-SNPs) were generated using Parsnp (version 1.2) software [44] and were visualized using Interactive Tree Of Life (iTOL v.5) [45]. The draft genomes of 287 human commensal *E. coli* (PRJNA400107) were downloaded from NCBI and included for phylogenetical analysis in this study [17].

Plasmid conjugation experiments

Plasmid conjugation experiments between carbapenem-resistant *E. coli*, colistin-resistant *E. coli*, and/or tigecycline-resistant *E. coli* (donors) and rifampin-resistant *E. coli* C600 (recipient) were performed as described previously [25]. Briefly, bacterial donor and recipient strains at mid-log phase (OD_{600} = 0.5 ~ 0.6) were mixed at a ratio of 1:3 (*v*/*v*). Bacterial mixture was spotted on nitrocellulose membranes that were pre-plated on LB agars. An incubation at 37°C for 12 h was given to each of the plates, and bacteria on the membrane were washed using LB broth followed by being shaken at 37°C for 4 h. Transconjugants were selected on LB agar plates with rifampin (1000 mg/L) plus imipenem (20 mg/L) [to screen carbapenem-resistant transconjugants], or rifampin (1000 mg/L) plus colistin (2 mg/L) [to screen colistin-

resistant transconjugants], or rifampin (1000 mg/L) plus tigecycline (4 mg/L) [to screen tigecycline-resistant transconjugants]. Antimicrobial susceptibility of the transconjugants was determined using broth microdilution method as mentioned above.

Declarations

Supporting Information

All authors declare no competing interests.

Acknowledgements

The authors sincerely acknowledge our colleagues for sample collection. Our work was supported by the National Key R&D Program of China (grant numbers: 2017YFC1600101 and 2017YFC1600103), the earmarked fund for China Agriculture Research System (grant number: CARS-35), the Natural Science Foundation of Hubei Province (grant number: 2020CFB525), China Postdoctoral Science Foundation (grant number: 2018M640719), Walmart Foundation (Project number: 61626817) and Walmart Food Safety Collaboration Center. The funders had no role in the study design, data collection, data analysis, data interpretation, or the manuscript writing.

References

1. CDC. Antibiotic resistance threats in the United States 2019. (2019).
2. Yao, X., Doi, Y., Zeng, L., Lv, L. & Liu, J.H. Carbapenem-resistant and colistin-resistant Escherichia coli co-producing NDM-9 and MCR-1. *Lancet Infect Dis* **16**, 288-9 (2016).
3. He, T. *et al.* Emergence of plasmid-mediated high-level tigecycline resistance genes in animals and humans. *Nat Microbiol* **4**, 1450-1456 (2019).
4. 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**, 161-8 (2016).
5. WHO. Antimicrobial resistance: global report on surveillance 2014. Geneva: Wolrd Health Organization, 2014. .
6. WHO. Worldwide country situation analysis: response to antimicrobial resistance. Geneva: Wolrd Health Organization, 2015.
7. Poirel, L. *et al.* Antimicrobial Resistance in Escherichia coli. *Microbiol Spectr* **6**(2018).
8. Shrestha, P. *et al.* Enumerating the economic cost of antimicrobial resistance per antibiotic consumed to inform the evaluation of interventions affecting their use. *Antimicrob Resist Infect Control* **7**, 98 (2018).
9. Kumarasamy, K.K. *et al.* Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study. *Lancet Infect Dis* **10**, 597-602 (2010).

10. Van Boeckel, T.P. *et al.* Reducing antimicrobial use in food animals. *Science* **357**, 1350-1352 (2017).
11. Van Boeckel, T.P. *et al.* Global trends in antimicrobial resistance in animals in low- and middle-income countries. *Science* **365**(2019).
12. Yang, H. *et al.* Antibiotic Application and Resistance in Swine Production in China: Current Situation and Future Perspectives. *Front Vet Sci* **6**, 136 (2019).
13. Zhang, Q.Q., Ying, G.G., Pan, C.G., Liu, Y.S. & Zhao, J.L. Comprehensive evaluation of antibiotics emission and fate in the river basins of China: source analysis, multimedia modeling, and linkage to bacterial resistance. *Environ Sci Technol* **49**, 6772-82 (2015).
14. Brisola, M.C. *et al.* Escherichia coli used as a biomarker of antimicrobial resistance in pig farms of Southern Brazil. *Sci Total Environ* **647**, 362-368 (2019).
15. Kaper, J.B., Nataro, J.P. & Mobley, H.L. Pathogenic Escherichia coli. *Nat Rev Microbiol* **2**, 123-40 (2004).
16. Chen, C. *et al.* Complete sequence of a tet(X4)-harboring IncX1 plasmid, pYY76-1-2, in Escherichia coli from a cattle sample in China. *Antimicrob Agents Chemother* **63**(2019).
17. Sun, C. *et al.* Genomic epidemiology of animal-derived tigecycline-resistant Escherichia coli across China reveals recent endemic plasmid-encoded tet(X4) gene. *Commun Biol* **3**, 412 (2020).
18. Li, J. Current status and prospects for in-feed antibiotics in the different stages of pork production - A review. *Asian-Australas J Anim Sci* **30**, 1667-1673 (2017).
19. He, Y., Yang, X., Xia, J., Zhao, L. & Yang, Y. Consumption of meat and dairy products in China: a review. *Proc Nutr Soc* **75**, 385-91 (2016).
20. Li, Y.W. *et al.* [Investigation of sulfonamides and tetracyclines antibiotics in soils from various vegetable fields]. *Huan Jing Ke Xue* **30**, 1762-6 (2009).
21. Krishnasamy, V., Otte, J. & Silbergeld, E. Antimicrobial use in Chinese swine and broiler poultry production. *Antimicrob Resist Infect Control* **4**, 17 (2015).
22. Sun, J. *et al.* Co-transfer of bla(NDM-5) and mcr-1 by an IncX3-X4 hybrid plasmid in Escherichia coli. *Nat Microbiol* **1**, 16176 (2016).
23. Zhai, R. *et al.* Contaminated in-house environment contributes to the persistence and transmission of NDM-producing bacteria in a Chinese poultry farm. *Environ Int* **139**, 105715 (2020).
24. Wang, Y. *et al.* Changes in colistin resistance and mcr-1 abundance in Escherichia coli of animal and human origins following the ban of colistin-positive additives in China: an epidemiological comparative study. *Lancet Infect Dis* **20**, 1161-1171 (2020).
25. Peng, Z. *et al.* Characteristics of Carbapenem-Resistant and Colistin-Resistant Escherichia coli Co-Producing NDM-1 and MCR-1 from Pig Farms in China. *Microorganisms* **7**(2019).
26. Foong, W.E., Wilhelm, J., Tam, H.K. & Pos, K.M. Tigecycline efflux in Acinetobacter baumannii is mediated by TetA in synergy with RND-type efflux transporters. *J Antimicrob Chemother* **75**, 1135-1139 (2020).

27. Croxen, M.A. *et al.* Recent advances in understanding enteric pathogenic Escherichia coli. *Clin Microbiol Rev* **26**, 822-80 (2013).
28. Liu, D. Diarrhoeagenic Escherichia coli. *Molecular Medical Microbiology (Second Edition)* **2**, 1127-1131 (2015).
29. Johnson, T.J. & Nolan, L.K. Pathogenomics of the virulence plasmids of Escherichia coli. *Microbiol Mol Biol Rev* **73**, 750-74 (2009).
30. Li, X. *et al.* Dissemination of bla(NDM-5) gene via an IncX3-type plasmid among non-clonal Escherichia coli in China. *Antimicrob Resist Infect Control* **7**, 59 (2018).
31. Sun, J. *et al.* Plasmid-encoded tet(X) genes that confer high-level tigecycline resistance in Escherichia coli. *Nat Microbiol* **4**, 1457-1464 (2019).
32. Sabat, G., Rose, P., Hickey, W.J. & Harkin, J.M. Selective and sensitive method for PCR amplification of Escherichia coli 16S rRNA genes in soil. *Appl Environ Microbiol* **66**, 844-9 (2000).
33. Wirth, T. *et al.* Sex and virulence in Escherichia coli: an evolutionary perspective. *Mol Microbiol* **60**, 1136-51 (2006).
34. Peng, Z. *et al.* Genetic and Phylogenetic Characteristics of Pasteurella multocida Isolates From Different Host Species. *Front Microbiol* **9**, 1408 (2018).
35. Zankari, E. *et al.* Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* **67**, 2640-4 (2012).
36. Carattoli, A. *et al.* In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* **58**, 3895-903 (2014).
37. Joensen, K.G., Tetzschner, A.M., Iguchi, A., Aarestrup, F.M. & Scheutz, F. Rapid and Easy In Silico Serotyping of Escherichia coli Isolates by Use of Whole-Genome Sequencing Data. *J Clin Microbiol* **53**, 2410-26 (2015).
38. Larsen, M.V. *et al.* Multilocus sequence typing of total-genome-sequenced bacteria. *J Clin Microbiol* **50**, 1355-61 (2012).
39. Nakamura, T., Yamada, K.D., Tomii, K. & Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **34**, 2490-2492 (2018).
40. Aziz, R.K. *et al.* The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* **9**, 75 (2008).
41. Goris, J. *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**, 81-91 (2007).
42. Alikhan, N.F., Petty, N.K., Ben Zakour, N.L. & Beatson, S.A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402 (2011).
43. Sullivan, M.J., Petty, N.K. & Beatson, S.A. Easyfig: a genome comparison visualizer. *Bioinformatics* **27**, 1009-10 (2011).
44. Treangen, T.J., Ondov, B.D., Koren, S. & Phillippy, A.M. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol* **15**, 524

(2014).

45. Letunic, I. & Bork, P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* **44**, W242-5 (2016).

Figures

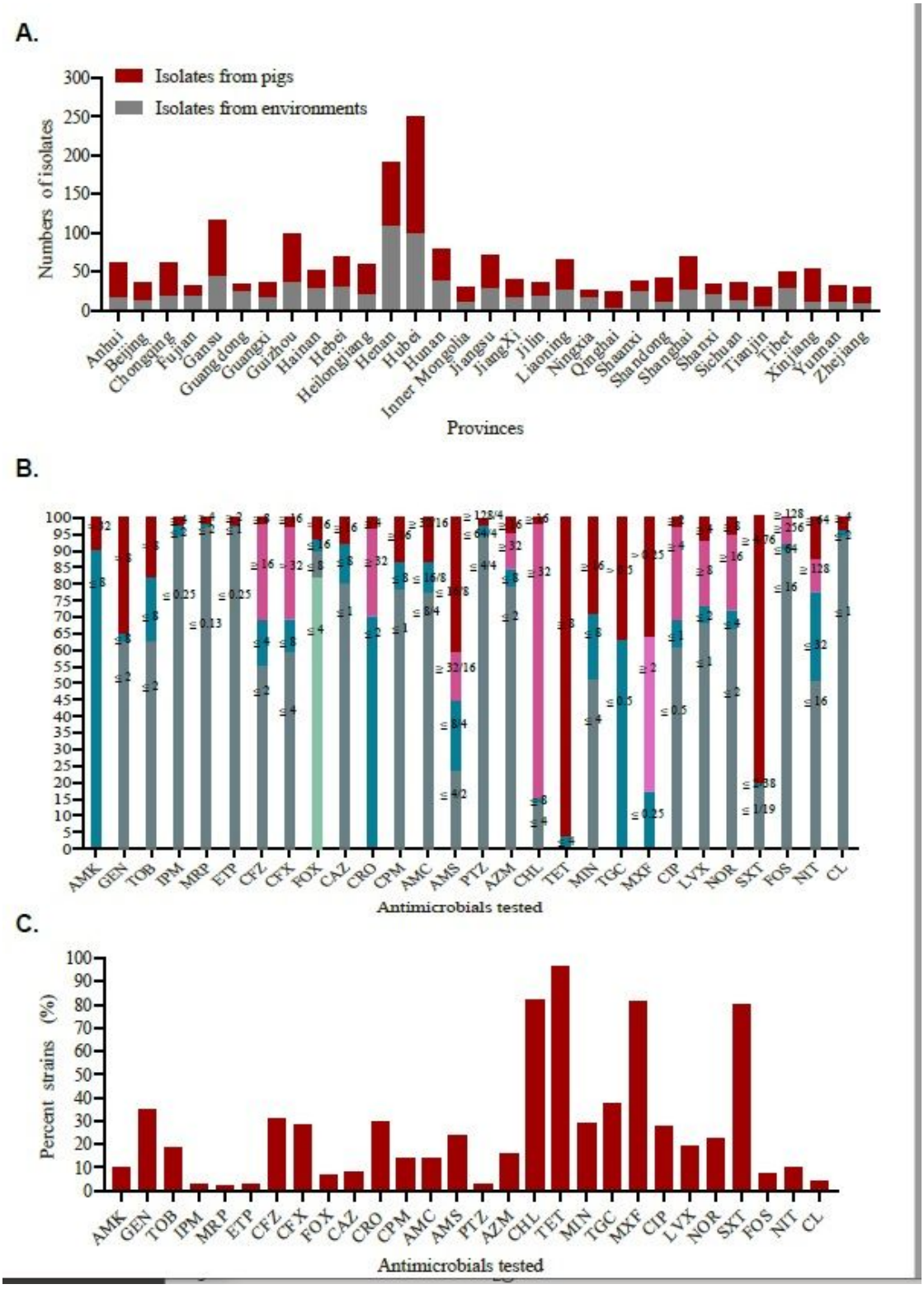

Figure 1

solation and antimicrobial resistance phenotypes of E. coli isolates from pig farms in different provinces in China. (A.) Column chart showing the number of E. coli isolates from pig associated samples and environmental samples collected from pig farms in each of provinces in China; (B.) Column chart showing the distribution of minimum inhibitory concentration (MIC) values of each of the tested antibiotics on E. coli isolates from pig farms in China; (C.) Column chart showing the percentage of farm-origin E. coli isolates resisting different antibiotics tested. AMK: amikacin, GEN: gentamicin, TOB: tobramycin, IPM: imipenem, MRP: meropenem, ETP: ertapenem, CFZ: cefazolin, CFX: cefuroxime, FOX: cefoxitin, CAZ: ceftazidime, CRO: ceftriaxone, CPM: cefepime, AMC: amoxicillin/clavulanate, AMS: ampicillin/sulbactam, PTZ: piperacillin/tazobactam, AZM: aztreonam, CHL: chloramphenicol, TET: tetracycline, MIN: minocycline, TGC: tigecycline, MXF: moxifloxacin, CIP: ciprofloxacin, LVX: levofloxacin, NOR: norfloxacin, SXT: trimethoprim/sulfamethoxazole, FOS: fosfomycin, NIT: nitrofurantoin, CL: colistin.

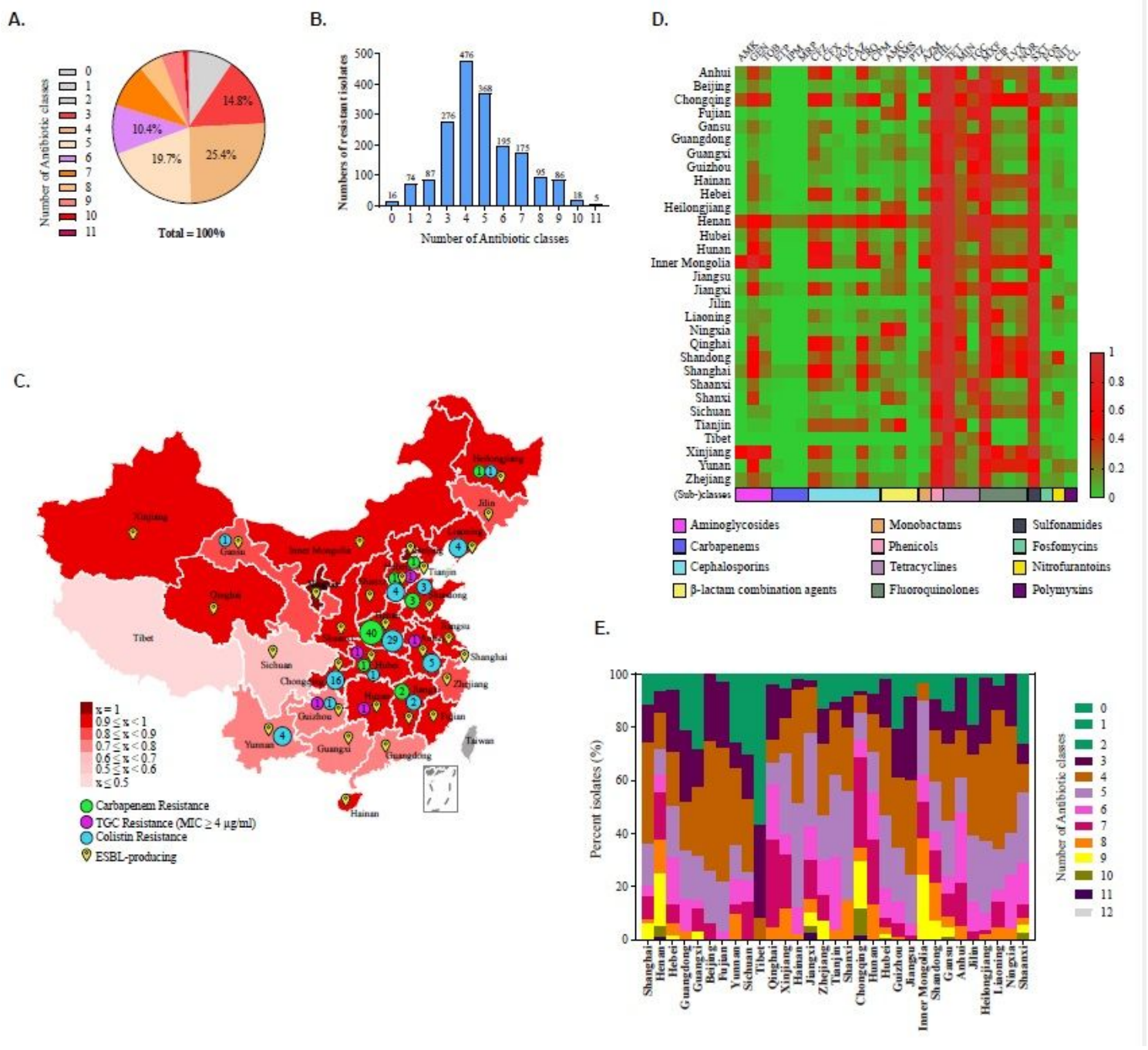

Figure 2

Antimicrobial resistance characteristics of E. coli isolates from pig farms in different provinces in China. (A.) Percentages of E. coli isolates from pig farms in different provinces in China resisting different antibiotic classes tested; (B.) Numbers of E. coli isolates from pig farms in different provinces in China resisting different antibiotic classes tested; (C.) A China map showing percentages of multidrug resistant E. coli isolates from pig farms in different provinces in China ("x" refers to percent MDR isolates); different colors representing the percentage; small circles in green, purple and blue show the distribution of carbapenem-resistant, colistin-resistant, and tigecycline-resistant E. coli isolates; numbers in small circles represent the numbers of carbapenem-resistant, colistin-resistant, and tigecycline-resistant E. coli isolates in different provinces; ESBL-producing E. coli isolates are also marked; (D.) Heat-map showing

percent farm-origin E. coli isolates resisting each of the antibiotics tested in different provinces in China; (E.) Percentages of farm-origin E. coli isolates resisting different antibiotic classes in different provinces in China. AMK: amikacin, GEN: gentamicin, TOB: tobramycin, IPM: imipenem, MRP: meropenem, ETP: ertapenem, CFZ: cefazolin, CFX: cefuroxime, FOX: cefoxitin, CAZ: ceftazidime, CRO: ceftriaxone, CPM: cefepime, AMC: amoxicillin/clavulanate, AMS: ampicillin/sulbactam, PTZ: piperacillin/tazobactam, AZM: aztreonam, CHL: chloramphenicol, TET: tetracycline, MIN: minocycline, TGC: tigecycline, MXF: moxifloxacin, CIP: ciprofloxacin, LVX: levofloxacin, NOR: norfloxacin, SXT: trimethoprim/sulfamethoxazole, FOS: Fosfomycin, NIT: nitrofurantoin, CL: colistin. Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

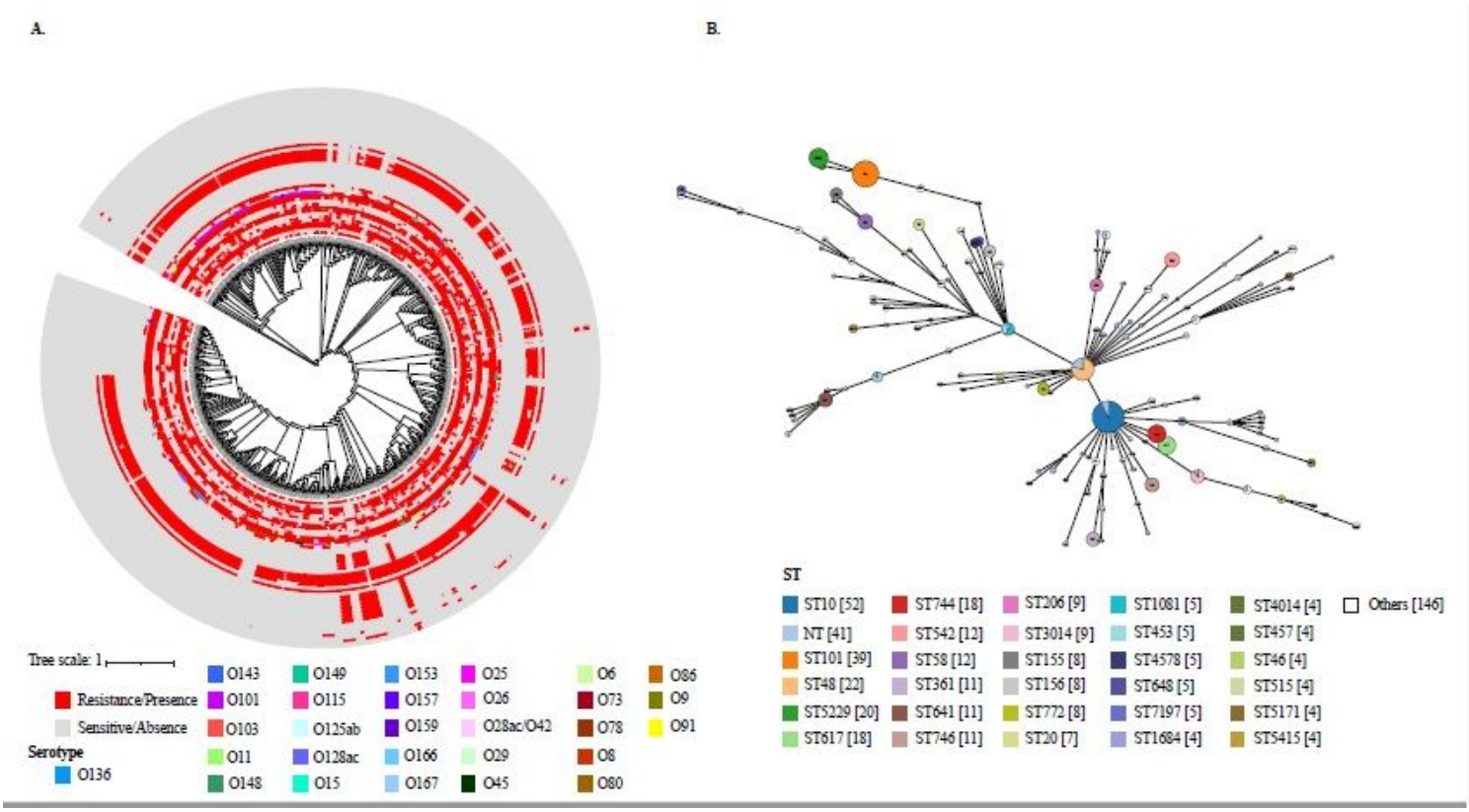

Figure 3

Population genomics of extensively drug resistant E. coli isolates from pig farms in different provinces in China. (A.) Phylogenetic analysis of the 515 sequenced extensively drug resistant E. coli isolates from pig farms in China; Circles from inside to outside: resistance phenotypes of E. coli isolates to aminoglycosides (circle 1), carbapenems (circle 2), 1/2-generation-cephalosporins (circle 3), 3/4-generation-cephalosporins (circle 4), monobactams (circle 5), β-lactam combination agents (circle 6), colistin (circle 7), sulfonamides (circle 8), phenicols (circle 9), fosfomycins (circle 10), nitrofurantoins (circle 11), fluoroquinolones (circle 12), and tetracyclines (circle 13); O-serogroups of the isolates (circle 14); presence of virulence factors encoding genes astA (circle 15), eae (circle 16), east1 (circle 17), ecpA

(circle 18), ecpB (circle 19), ecpC (circle 20), ecpD (circle 21), ecpE (circle 22), ecpR (circle 23), efa1 (circle 24), eltA (circle 25), eltB (circle 26), escC (circle 27), escD (circle 28), escF (circle 29), escJ (circle 30), escN (circle 31), escR (circle 32), escS (circle 33), escT (circle 34), escU (circle 35), escV (circle 36), espA (circle 37), espB (circle 38), espD (circle 39), estIa (circle 40), paa (circle 41), pic (circle 42), stb (circle 43), stx2eB (circle 44), tir (circle 45), and toxB (circle 46); (B.) Analysis of Minimum Spanning Tree of the 515 sequenced extensively drug resistant E. coli isolates based on the sequence type.

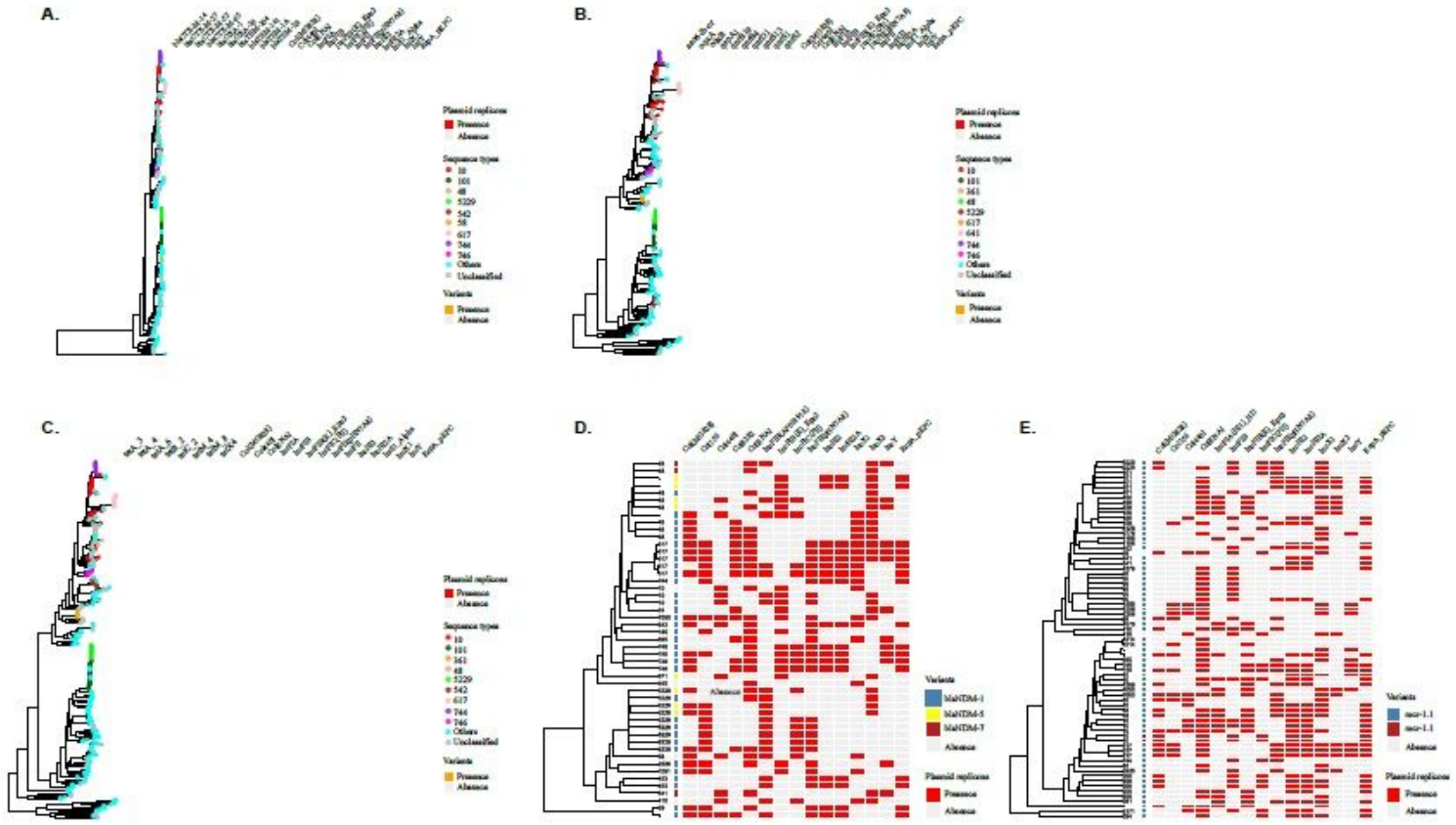

Figure 4

Genomic characteristics of farm origin E. coli isolates with resistance phenotypes to specific (sub)classes of antibiotics. (A.) Phylogenetic analysis of extended-spectrum beta-lactamase (ESBL)-producing E. coli isolates from pig farms in China; sequence types, presence of ESBL-genes, as well as determined main plasmid types are shown in the figure; (B.) Phylogenetic analysis of fluoroquinolone-resistant E. coli isolates from pig farms in China; sequence types, presence of fluoroquinolone-resistant genes, as well as determined main plasmid types are shown in the figure; (C.) Phylogenetic analysis of tetracycline-resistant E. coli isolates from pig farms in China; sequence types, presence of tetracycline-resistant genes, as well as determined main plasmid types are shown in the figure; (D.) Phylogenetic analysis of carbapenem-resistant E. coli isolates from pig farms in China; sequence types, presence of carbapenem-resistant genes, as well as determined main plasmid types are shown in the figure; (E.) Phylogenetic analysis of colistin-resistant E. coli isolates from pig farms in China; sequence types, presence of colistin-resistant genes, as well as determined main plasmid types are shown in the figure.

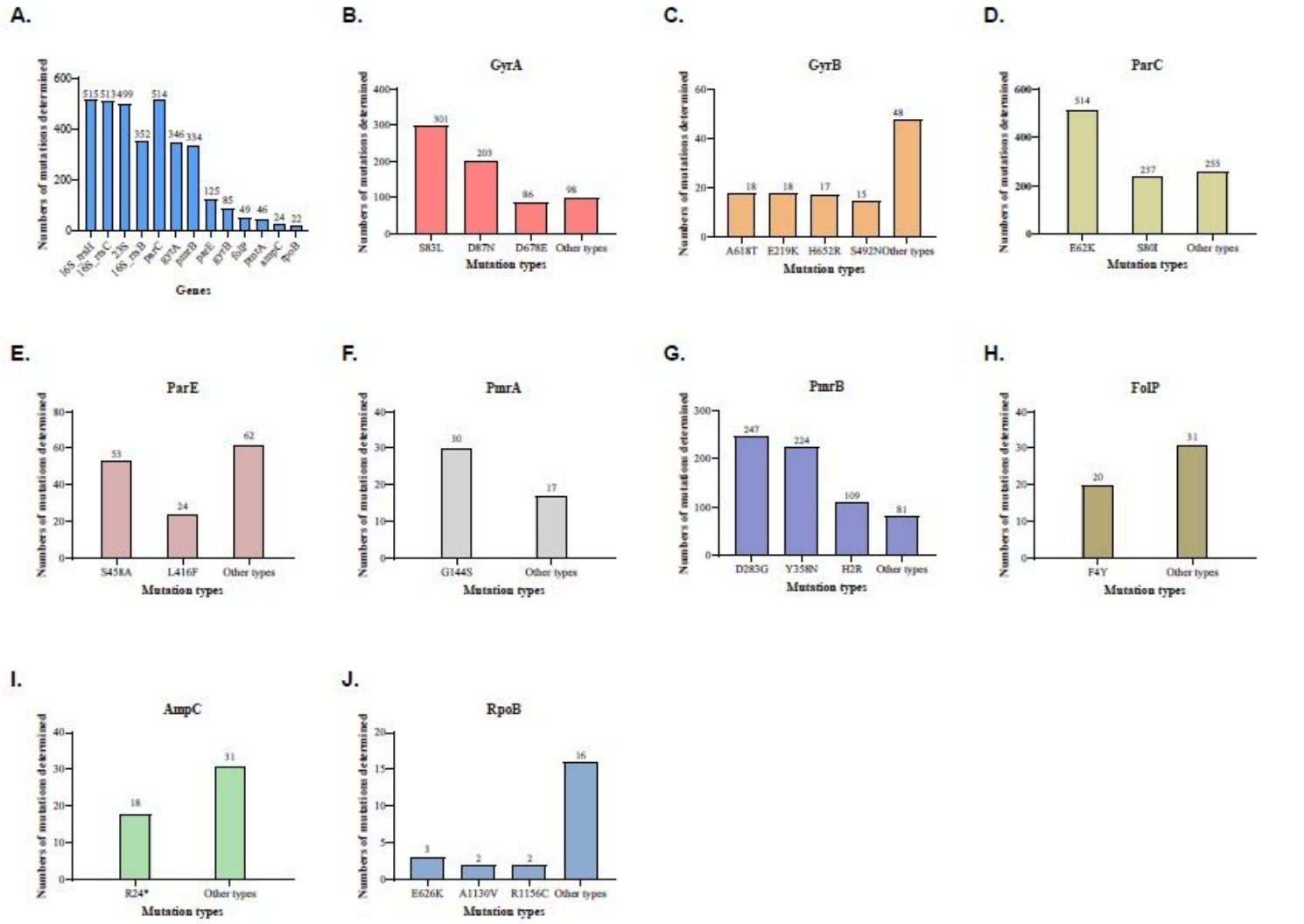

Figure 5

Point mutations determined in extensively drug resistant E. coli isolates from pig farms in different provinces in China. Panel A shows the number of E. coli isolates that carry genes with antimicrobial resistance associated point mutations; Panels B~J exhibit the distribution of different mutation types determined in GyrA (panel B), GyrB (panel C), ParC (panel D), ParE (panel E), PmrA (panel F), PmrB (panel G), FolP (panel H), AmpC (panel I), and RpoB (panel J).

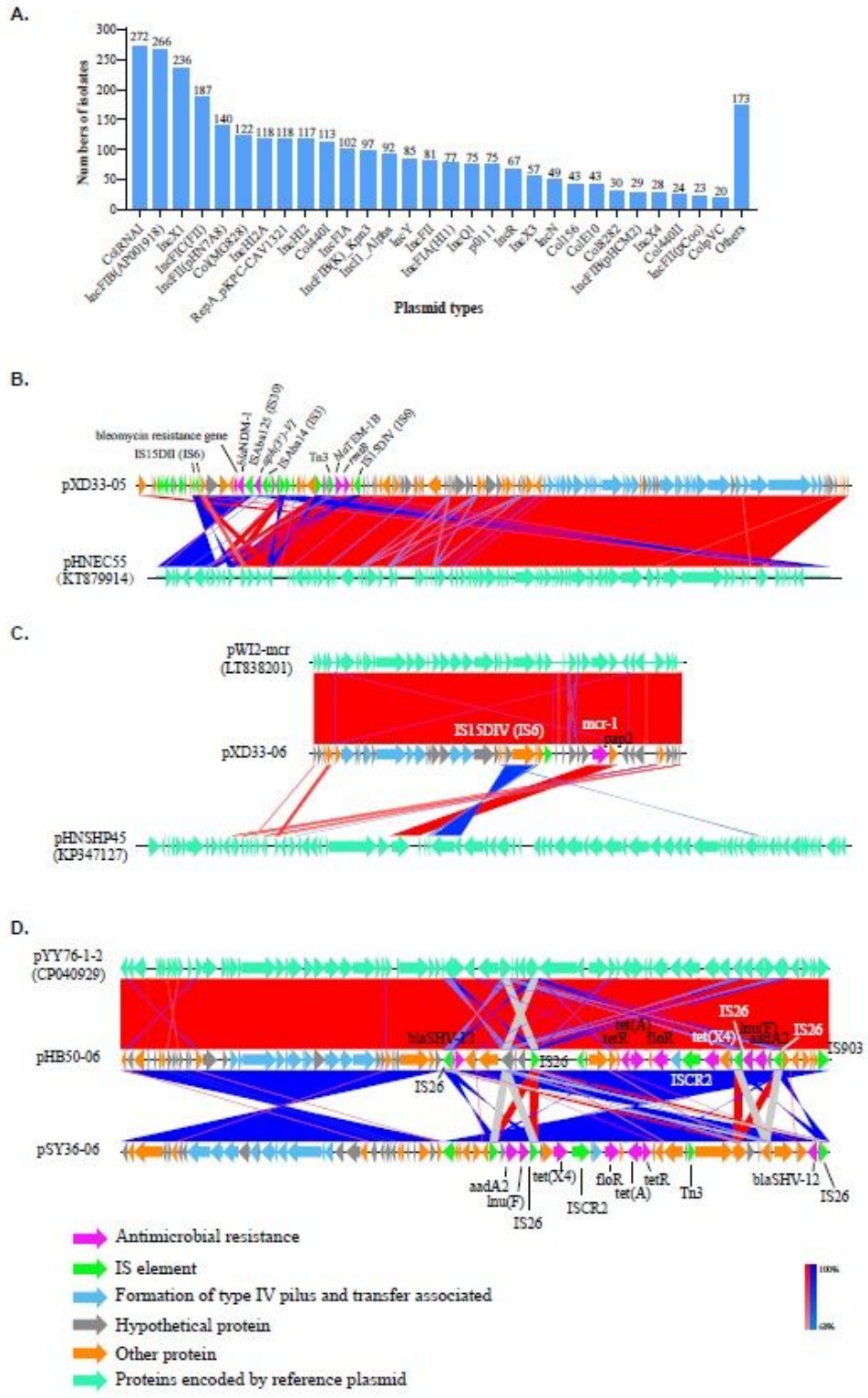

Figure 6

Plasmids mediating the dissemination of blaNDM-1, mcr-1, and tet(X4) determined in selected E. coli isolates from pig farms in China. Panel A shows the distribution of different plasmid types determined in farm originated extensively drug resistant E. coli isolates in China; Panels B~D exhibit genetic structures of blaNDM-1-bearing plasmid (panel B), mcr-1-bearing plasmid (panel C), and tet(X4)-bearing plasmid

(panel D). Color code stands for BLASTn identity of those regions between genomes. Arrows stand for putative CDSs in different genomes.

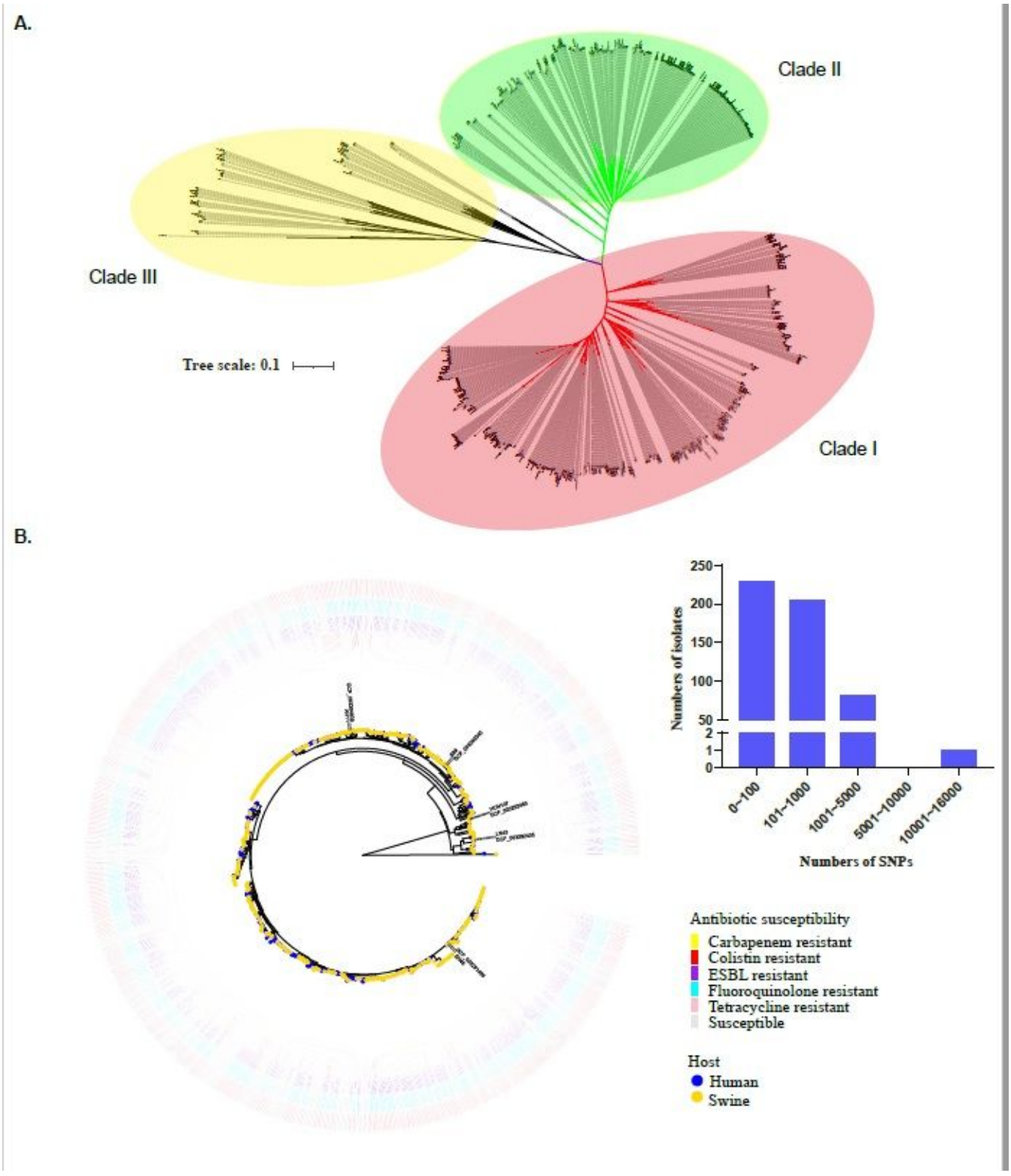

Figure 7

Phylogenetic trees of 515 XDR-E. coli isolates from pig farms together with 287 publicly available draft genomes of human commensal E. coli (Bioproject no. PRJNA4001047) across China. (A.) An unrooted tree showing the three Bayesian lineages. (B.) A circular tree showing the corresponding information and

SNP differences between E. coli isolates from humans and/or pig farms. Five pairs of farm-originated E. coli isolates and human isolates that shared less than 10 SNPs are displayed. Numbers of SNPs between the farm originated XDR-E. coli isolates and the human isolates are also shown in a column chart.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- TableS1.xlsx
- TableS2.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- FigureS1.tif
- FigureS2.tif
- nrreportingsummaryNCOMMS2117708.pdf