

RESEARCH ARTICLE

Open Access



Genomic characterization of an emerging Enterobacteriaceae species: the first case of co-infection with a typical pathogen in a human patient

Zhao Zhang^{1,2†}, Daixi Li^{1,3†}, Xing Shi^{1,4†}, Yao Zhai⁵, Yatao Guo^{1,2}, Yali Zheng^{1,6}, Lili Zhao¹, Yukun He¹, Yusheng Chen⁷, Zhanwei Wang⁸, Jianrong Su⁹, Yu Kang^{4*} and Zhancheng Gao^{1*}

Abstract

Background: Opportunistic pathogens are important for clinical practice as they often cause antibiotic-resistant infections. However, little is documented for many emerging opportunistic pathogens and their biological characteristics. Here, we isolated a strain of extended-spectrum β -lactamase-producing *Enterobacteriaceae* from a patient with a biliary tract infection. We explored the biological and genomic characteristics of this strain to provide new evidence and detailed information for opportunistic pathogens about the co-infection they may cause.

Results: The isolate grew very slowly but conferred strong protection for the co-infected cephalosporin-sensitive *Klebsiella pneumoniae*. As the initial laboratory testing failed to identify the taxonomy of the strain, great perplexity was caused in the etiological diagnosis and anti-infection treatment for the patient. Rigorous sequencing efforts achieved the complete genome sequence of the isolate which we designated as AF18. AF18 is phylogenetically close to a few strains isolated from soil, clinical sewage, and patients, forming a novel species together, while the taxonomic nomenclature of which is still under discussion. And this is the first report of human infection of this novel species. Like its relatives, AF18 harbors many genes related to cell mobility, various genes adaptive to both the natural environment and animal host, over 30 mobile genetic elements, and a plasmid bearing *bla*_{CTX-M-3} gene, indicating its ability to disseminate antimicrobial-resistant genes from the natural environment to patients. Transcriptome sequencing identified two sRNAs that critically regulate the growth rate of AF18, which could serve as targets for novel antimicrobial strategies.

Conclusions: Our findings imply that AF18 and its species are not only infection-relevant but also potential disseminators of antibiotic resistance genes, which highlights the need for continuous monitoring for this novel species and efforts to develop treatment strategies.

Keywords: Enterobacteriaceae, Pathogen, Whole-genome sequencing, RNA-Seq, Phylogenetic

* Correspondence: kangy@big.ac.cn; zcgao@bjmu.edu.cn

[†]Zhao Zhang, Daixi Li and Xing Shi contributed equally to this work.

⁴Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, Beijing, China

¹Department of Respiratory & Critical Care Medicine, Peking University People's Hospital, Beijing, Beijing, China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Antimicrobial resistance (AMR) is an increasingly global health threat that contributes to 700,000 deaths per year [1]. Increased and often unrestricted antibiotic use in the clinical and farming settings is to blame for this issue. Growing surveillances based on genomic sequencing of microbes from the natural environment, human settlements, and clinical settings have been conducted worldwide to investigate the evolution and transfer of antibiotic resistance genes (ARGs) [2–4]. In recent years, the eco-evolutionary feedback loops between ecological and evolutionary dynamics have been increasingly recognized, where spillover of antibiotic use to natural and semi-natural environments may have profound implications on the distribution of ARGs in natural bacterial populations which serve as environmental reservoirs of resistance determinants [5, 6]. However, how resistance evolves, and how ARGs are maintained and dispersed back to clinical settings is poorly understood. Understanding the dynamics of the continuous feedback loops from clinical to nature and back may prove critical for preventing and controlling the problem of antibiotic resistance.

The rapidly developing sequencing technology increasingly enables the identification of emerging opportunistic pathogens and taxonomical classification based on their genomic information [7–9]. Naturally, opportunistic pathogens inhabit in the natural environment and are occasionally resistant to common antibiotics. Among these previously unknown pathogens, many are belong to species of the *Enterobacteriaceae* family [10, 11]. Meanwhile, many *Enterobacteriaceae* species are commensal microbiota of human and animal guts, but under certain conditions, can be opportunistic pathogens that cause infections [12]. These species often have other animal hosts, or they can be found in more diverse environments, such as soil and sewage [13]. *Enterobacteriaceae* species (including *E. coli*, *Klebsiella*, and *Enterobacter*) are also famous for their antibiotic resistance and regarded as some of the most dangerous pathogens since they can efficiently acquire various ARGs through efficient plasmid transmission [14]. The ability of these species to disseminate between habitats and transferring ARGs highlights their importance as mediators in the eco-evolutionary feedback loops that disperse ARGs from natural environments back to clinical settings. The taxonomy of *Enterobacteriaceae* is complex, containing 28 genera and over 75 species [15], while novel species are continuously discovered. Recognizing and characterizing *Enterobacteriaceae* species, especially those of emerging opportunistic pathogens, is critical for understanding the dynamics of the evolution of AMR.

Here, we isolated from a patient with a biliary infection a novel strain of unknown taxonomy accompanying an infectious *Klebsiella pneumoniae* strain, which we

designated as AF18. AF18 grew slowly but provided drug-resistance to its companion by carrying a *bla*_{CTX-M-3} resistant gene. The co-infection brought perplexity in both diagnosis and treatment of the patients. In order to provide new evidence and detailed information for opportunistic pathogens about the complex issues that they may cause in clinical infections, we conducted a study with the three following objectives: (1) Clarifying the taxonomy of AF18 using whole-genome phylogenetic analysis; (2) Testing the ability of AF18 to protect *K. pneumoniae* from antibiotics in co-culture experiments; and (3) Analyzing the adaptation mechanisms of AF18 base on transcriptome sequencing. Finally, we find that AF18 is a strain of an undefined novel species in the family *Enterobacteriaceae*, and that sensitive *K. pneumoniae* can survive when co-cultured with AF18 in Luria-Bertani broth containing 8 µg/mL ceftriaxone. Furthermore, genomic and transcriptomic analyses reveal the genomic characteristics of this rare pathogen and the regulation mechanisms of how it adapts to multiple habitats and its association with ARGs transfer.

Results

Biological identification of the strain AF18

From the bile sample of the patient, two types of colonies were isolated after serial dilutions and isolations on MacConkey agar plates. One type was mucous, entirely pink, and of 4–5 mm in diameter, which was finally identified as a *K. pneumoniae* clone sensitive to common antibiotics (Table 1); the other type was composed by small (2–3 mm in diameter) red-centered colonies with clear and transparent edges (Fig. 1a). The bacteria of the small colonies seemed prone to adhere to the cells of *K. pneumoniae* and were not able to be isolated until extensive dilutions. The taxonomy of the small colonies was not immediately identified by the microbiological laboratory in the hospital, and we designated it as strain AF18. AF18 exhibited resistance to most β-lactam antibiotics in antimicrobial susceptibility testing (Table 1). As the infection was rather intractable and finally cured by intravenous amikacin, the final diagnosis for the patient was a co-infection caused by a sensitive *K. pneumoniae* strain and a multidrug-resistant strain of unknown species.

Microscope observation showed that AF18 was a Gram-negative bacillus (Fig. 1b), and its cells were surrounded by flagella under a transmission electron microscope (Fig. 1c). The scanning electron microscope confirmed the tubular shape of AF18 and a smooth surface with no polysaccharide particles (Fig. 1d), in line with the mucus-free characteristics of its colonies. VITEK-II in the hospital laboratory did not identify any bacterial species with identical biochemical properties to AF18 (Table S1), whereas the API20E biochemical

Table 1 The antibiotic resistance profile of AF18 and *K. pneumoniae* isolate

Drug	Antibiotic susceptibility			
	AF18		<i>K. pneumoniae</i> strain	
	MIC ($\mu\text{g/ml}$)	Phenotype	MIC ($\mu\text{g/ml}$)	Phenotype
Ampicillin	≥ 32	R	16	I
Ampicillin/sulbactam	≥ 32	R	4	S
Piperacillin	≥ 128	R	≤ 4	S
Piperacillin/tazobactam	≥ 128	R	≤ 4	S
Cefazolin	≥ 64	R	≤ 4	S
Cefuroxime	≥ 64	R	≤ 1	S
Cefuroxime axetil	≥ 64	R	≤ 1	S
Cefotetan	≤ 4	S	≤ 4	S
Ceftazidime	16	R	≤ 1	S
Ceftriaxone	≥ 64	R	≤ 1	S
Cefepime	≥ 64	R	≤ 1	S
Aztreonam	≥ 64	R	≤ 1	S
Imipenem	≤ 1	S	≤ 1	S
Meropenem	≤ 0.25	S	≤ 0.25	S
Amikacin	≤ 2	S	≤ 2	S
Gentamicin	≤ 1	S	≤ 1	S
Tobramycin	2	S	≤ 1	S
Ciprofloxacin	2	I	≤ 0.25	S
Levofloxacin	1	S	≤ 0.25	S
Nitrofurantoin	256	R	≤ 16	S
Trimethoprim/sulfam	≤ 20	S	≤ 20	S

identification system suggested AF18 as *Pantoea* sp. but with low reliability. The mass spectrometry which scans the protein profile of samples did not identify the species of AF18 either.

Complete genome of Enterobacteriaceae bacterium AF18

To determine the taxonomy and genetic features of AF18, we performed whole-genome sequencing using two platforms, Illumina Hiseq (generates short-reads) and PacBio sequencer (generates long-reads), obtaining a high-quality completed genome sequence. AF18 possessed a circulated chromosome and two plasmids (Table 2).

By using Mash [16] to search the publicly available bacterial genomes and drafts with a cutoff of mutation distance < 0.25 , we identified 33 non-redundant close relatives of AF18, all of which were in the *Enterobacteriaceae* family (Table S2). The average nucleotide identity (ANI) matrix of the 34 strains (Fig. 2a) shows that the closest five with identity $> 98.5\%$ ($> 95\%$ regarded as strains of the same species [17]) are nominated as [*Kluyvera*] *intestini* (GCA_001856865.3), *Metakosakonia* sp.(GCA_003925915.1), *Enterobacter* sp. (GCA_000814915.1, GCA_900168315.1), and just *Enterobacteriaceae bacterium* (GCA_002903045.1). The phylogenetic relationship of these relatives was further

inferred with core genome SNPs (Fig. 2b), which confirmed the relationships inferred from the ANI matrix and indicated the novel species, including AF18, possibly represents another genus than *Kluyvera*. Herein, we temporarily nominated our stain as *Enterobacteriaceae bacterium* AF18 as the nomenclature of its genus and species is still undefined.

We predicted seven copies of 16S rDNA sequences in AF18. We aligned them to the 33 genomes we picked using BLASTN and calculated the average identity. We removed the genomes which do not contain high quality 16S rDNA sequence. The result shows a good congruence of 16S rDNA and whole-genome comparisons (Table S3). However, considering cutoffs commonly used for intra-species classification by whole-genome ANI $> 95\%$ [17] and 16S rDNA identity $> 99\%$ [18], 16S rDNA classification found two more strains of the species, namely *Enterobacteriaceae bacterium* ENNIH1, and *Phytobacter ursingii* strain CAV1151 (Table S3). Thus, we think that 16S rDNA can also be used as a marker gene to clarify the taxonomy of isolated strains, but we need to examine the identity cutoff we used carefully.

The chromosome of AF18 possesses 5651 protein-coding genes whose functions facilitate the survival and adaptation of AF18 in various habits (Table S4, Table

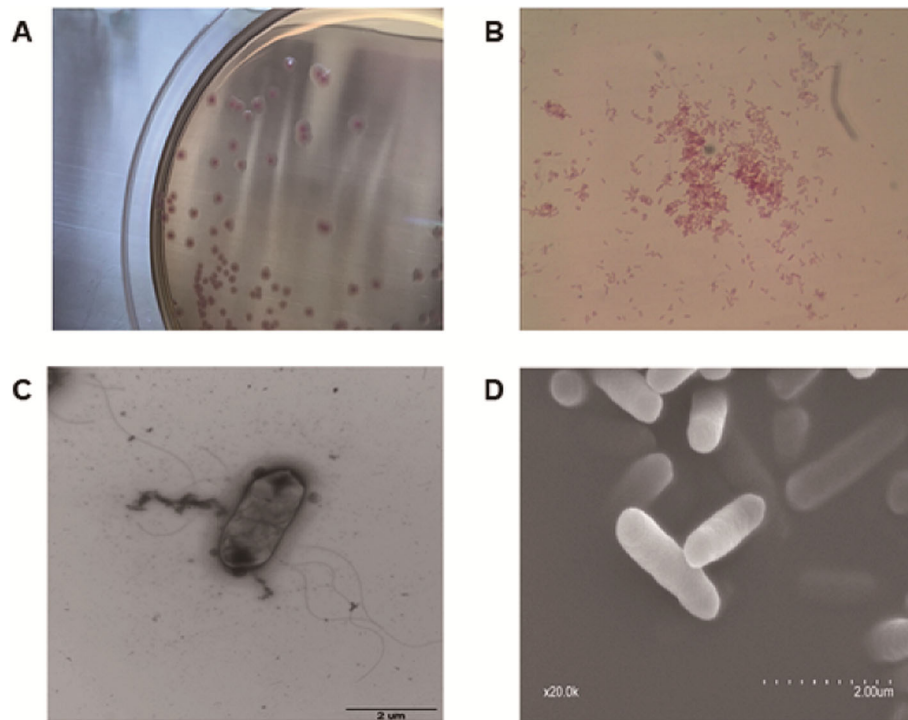


Fig. 1 The morphological characters of AF18. **a** The morphology of AF18 colonies on MacConkey agar plate. **b** Gram staining of AF18 cells. **c** Flagella of AF18 photographed by transmission electron microscopy. **d** Cells of AF18 under scanning electron microscopy

S5). For example, motility-related genes, including a complete flagellar gene cluster that encodes all components of flagellar, *csg* gene cluster that encodes curli assembly proteins to mediate adhesion, and other genes of *ompA*, *pilRT*, *ibeB*, *icaA*, *htpB* and *fimB*, together confer the ability of adhesion, invasion, chemotaxis, and escape to the host strain. Efflux pump genes which confer resistance to macrolides, quinolones and aminoglycosides were also identified. Meanwhile, the AF18 genome possesses 20 genomic islands, 11 prophages, and five CRISPR sequences (Table S5), suggesting the active transfer of stress-adaptive genes by these mobile genetic elements in this species. More importantly, markers of soil-inhabiting bacteria, including a complete nitrogen fixation gene cluster and *ksgA*— a pesticide-resistant gene, were found in AF18 genome, which suggests that AF18 is able to colonize natural environments. The

mobility of this strain may potentiate its dissemination to various habits.

Analysis of conserved genes in plasmids shows that most of the antibiotic-resistant genes of AF18, including *qnrS*, *dfrA*, and *bla*_{CTX-M-3}, are carried by the smaller plasmid pAF18_2 (Fig. 3, Table S4) which is, in major part, responsible for the antibiotic resistance profile of AF18 (Table 1). Sequence alignment shows that pAF18_2 is similar to many plasmids from other *Enterobacteriaceae* species, such as *E. coli* (KF914891.1, KC788405.1, CP028486.1), *K. pneumoniae* (KX928750.1, CP026179.1), and *C. freundii* (KT989599.1), and they contain identical replication origins, replication and transcription systems, plasmid partition systems, and a partial gene cluster responsible for plasmid conjugation, which indicates that the plasmid might be compatible with all these *Enterobacteriaceae* host species. Besides, these plasmids share a common anti-restriction system that ensures they would not be destroyed by the restriction-modified system in other host strains. Specifically, the pAF18_2 contains an active transposase system with complete IS elements which had acquired the *bla*_{CTX-M-3} gene and an arsenical resistant system. Many other DNA manipulating enzymes, such as integrase and DNA invertase, were also identified in the plasmid, all of which could facilitate the plasmid in efficiently acquiring and

Table 2 Overview of genome information for AF18

Replicon	Nucleotide length (bp)	Coding Genes	GC%	Inc type	GenBank ID
Chromosome	5,676,372	5651	53.06	NA	CP025982
pAF18_1	140,420	181	51.14	IncFII	CP025983
pAF18_2	42,923	53	51.28	IncN	CP025984

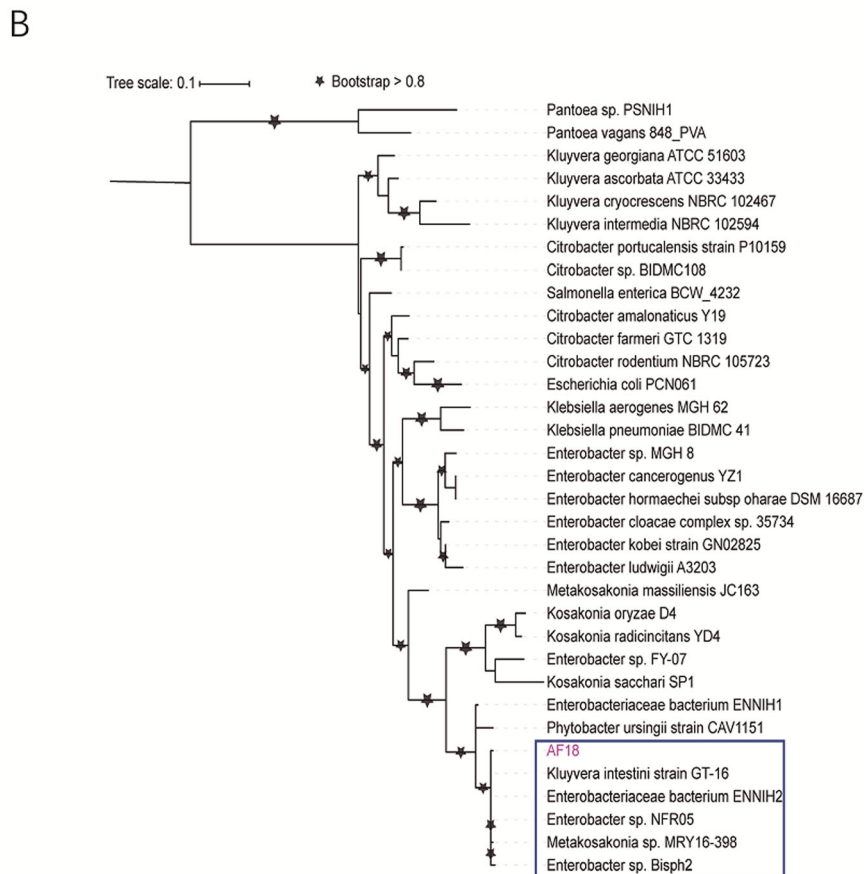
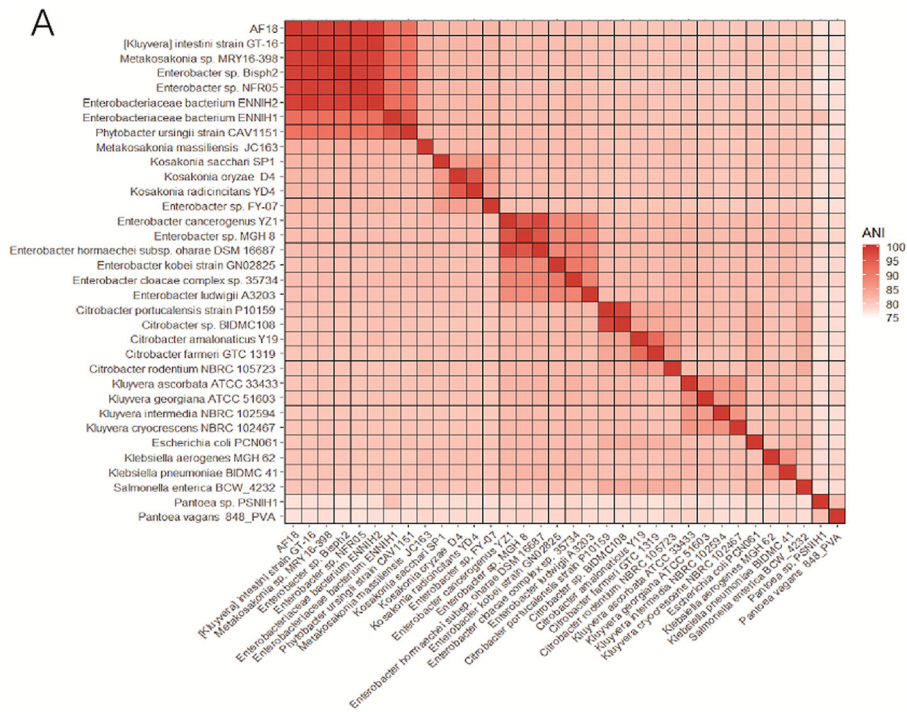


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Phylogenetic relationship of 34 strains related to AF18. **a.** The heatmap of ANI matrix. The color bar represents the value of ANI. The top five species (not including AF18) are the closest relatives of AF18 with ANI > 98.5%. **b.** The maximum likelihood phylogenetic tree constructed based on the core genome SNPs. The species in the blue box are the closest relatives of AF18 in the phylogenetic tree which are the same as the top five species of ANI heatmap. ANI, average nucleotide identity

transferring antibiotic-resistance genes and other stress-adaptive genes among *Enterobacteriaceae* strains. Unfortunately, due to constraints related to the outbreak of the 2019 novel coronavirus, we were unable to perform conjugation experiments.

Growth of AF18 in co-cultures and its transcriptional regulation

To disentangle the respective contribution of AF18 and the sensitive *K. pneumoniae* in the co-infection, we co-cultivated the two strain in various concentration of ceftriaxone, and found that addition of 1% of AF18 was

able to elevate the MIC from 0.125 µg/ml of pure *K. pneumoniae* culture to 64 µg/ml. Furthermore, when spreading the co-culture onto the MacConkey agar containing ceftriaxone, the sensitive *K. pneumoniae* colonies were able to withstand 8 µg/ml ceftriaxone (Fig. 4a), indicating a strong protective effect of AF18 to the co-infected *K. pneumoniae*.

Although necessary in the co-infection for antibiotic-resistance, AF18 only took less than 1% in the initial sample. Even when equally input, the proportion of AF18 decreased to 1% of the co-culture if without antibiotic pressure (Fig. 4b). It seems that AF18 may be less aggressive, and its growth rate is

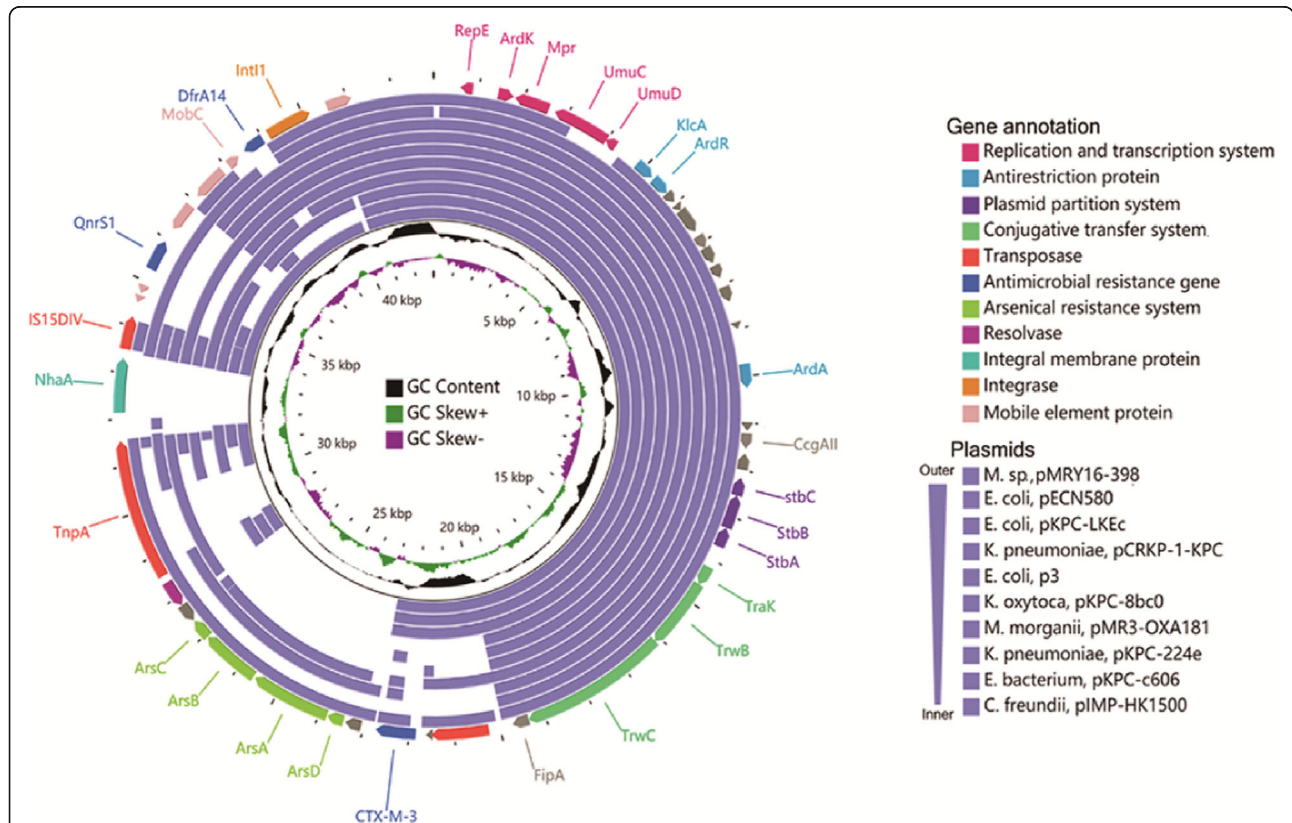
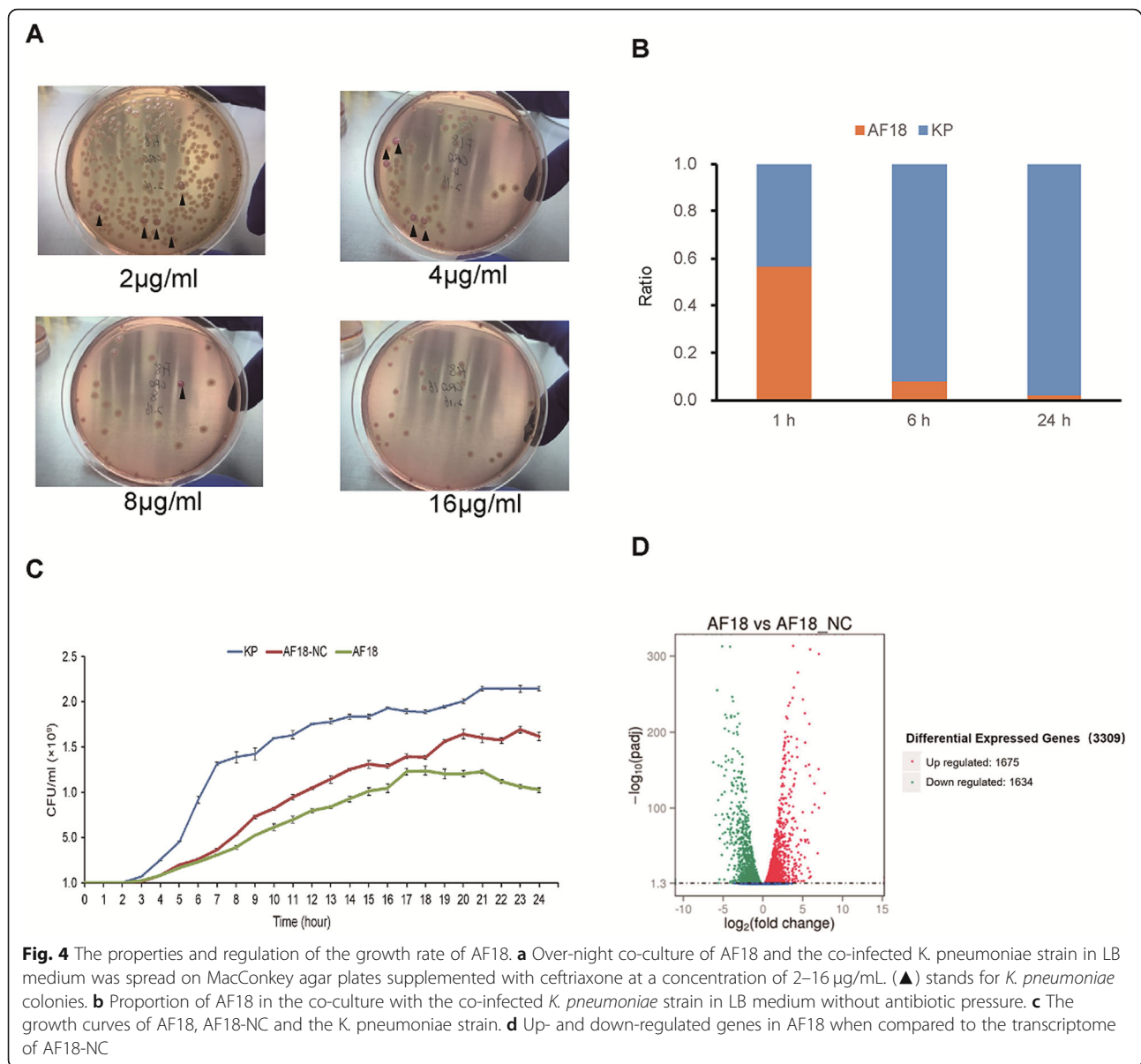


Fig. 3 The circular map of pAF18_2 and comparison to similar plasmids. The outmost slot represents the predicted genes of pAF18_2, whose functions are shown in different color arrows. From outward, slot 2–11 indicate aligned fragments from similar plasmids of IncN. Slot 12, GC content; slot 13, GC skew. Accession numbers of plasmids from outer to inner were: AP018758.1, KF914891.1, KC788405.1, KX928750.1, CP028486.1, CP026277.1, KM660724.1, CP026179.1, CP026198.1, KT989599.1



much slower than the co-inhabited *K. pneumoniae*. It has been reported that plasmid carriage may slow down growth rate due to the cellular cost imposed [19], and thus we generate a new strain—AF18-NC by deleting the resistant plasmid of AF18. Then we measured the independent growth curve of the three strains—*K. pneumoniae*, AF18, and AF18-NC, respectively (Fig. 4c). As expected, AF18-NC did grow faster than its mother strain AF18 since it was relieved from the plasmid-caused cellular cost. However, the growth rate of AF18-NC was still much slower than that of *K. pneumoniae*, suggesting that slow growth is an inherent property of the novel species.

Next, we analyzed the genes involved in the regulation of the growth rate by a comparison between the

transcriptomes of AF18 and AF18-NC. A total of 3309 genes of chromosomal coding genes were significantly differentially expressed, with 1675 upregulated and 1634 downregulated in AF18 (Fig. 4d). Functional cluster analysis with GO database showed that most of the differentially expressed genes were in the categories of transcriptional regulation, biosynthesis regulation, metabolic process regulation, signal transduction, and flagellar motility (Fig. S1). Analysis of the non-coding sRNA expression profile identified a total of 15 sRNAs differentially expressed between AF18 and AF18-NC. Interestingly, two of the down-regulated sRNAs in AF18, namely sRNA00063 and sRNA00291 (Fig. S2), shared 98% of their predicted target genes which constitute up to 56% of those differentially expressed genes as mentioned above,

suggesting that these two sRNAs have a key role in promoting growth. This result indicated the importance of the two sRNAs in global regulation of growth rate, and consequently, the contribution of the host AF18 in co-infections.

Discussion

In this study, we reported a case of co-infection caused by a typical pathogen and a rare opportunistic pathogen with taxonomical nomenclature undefined. *K. pneumoniae* is a common pathogen that can cause various aggressive infections [20], while the AF18 strain and its relatives have not been reported to cause infection in humans. In the pathogenic consortium of the co-infection in our study, although taking a very small proportion, AF18 provides strong protection for the entire pathogenic consortium against antibiotic damage. The cooperation between the *K. pneumoniae* and AF18 makes the infectious situation more complicated and difficult in terms of therapies than infections caused by either of them. Meanwhile, as the strain AF18 only took a minor proportion and closely adhered to co-infected *K. pneumoniae*, it was prone to be concealed by the dominant *K. pneumoniae* and hard to be detected and isolated, which led to inaccurate etiological diagnosis and improper anti-infective treatment at first admission. As AF18 and other strains of the same species are rare opportunistic pathogen with little documentation, and conventional testings for bacterial identification are not always correct for such novel species, as shown in this study, WGS comprised a straightforward approach for accurate taxonomy identification.

Sequencing strategy combining both long- and short-read platforms makes it easy to obtain high-quality complete genome and plasmids, which will be helpful for the overall characterization of novel species and deep insight into the functions of the genes they harbor. Phylogenetic analysis with the whole genome of AF18 assigned it to the Enterobacteriaceae family. However, the nomenclature of this novel species is still under discussion due to very limited documentation [21]. The first report of the novel species was in 2016 when [*Kluyvera*] *intestini* str. GT-16 was isolated from the stomach of a patient with gastric cancer [22], and in the following years, strains of this species were emergingly discovered [23, 24]. Of note, AF18 is the first clear report of human infection of this novel species, as [*Kluyvera*] *intestine* GT-16 and *Matakosakonia* sp. MRY16_398 were more likely a common resident in the gastrointestinal tract or a by-stander of the diverticulitis. Although the first strain of [*Kluyvera*] *intestini* str. GT-16 had been assigned to the genus *Kluyvera*, the ANI of strains in this novel species to typical *Kluyvera* spp. are less than 80.8%, even farther than the distance to other genera,

such as *Kasokonia* (ANI, 82.3%), and typical *Enterobacter* spp. (ANI, 81%), suggesting that AF18 and its species is not a typical *Kluyvera* species or should not be included in this genus. Except AF18 which cause a co-infection with a typical pathogen, two of the other strains of this species ([*Kluyvera*] *intestini* str. GT-16 [25], and *Matakosakonia* sp. MRY16_398 [23]) were both isolated from patients in Japan, while the other three strains (*Enterobacter* sp. NFR05, *Enterobacter* sp. Bisph2, and *Enterobacteriaceae* bacterium ENNIH2) were isolated from rhizoplane (China), sandy soil (Algeria) [26], and hospital sewage (USA) [24], respectively, indicating a wide range of environments that this species can inhabit.

16S rDNA sequencing is also a commonly used method for species identification; however, there are some matters to consider when using 16S rDNA to determine the taxonomy of a new bacteria. Briefly speaking, we need to obtain the entire 16S rDNA gene (preferably all copies of the strain) and an up-to-date database to ensure accurate identification [18]. Another issue that requires attention is that we indeed need some online websites and databases to help us identify bacterial taxonomy and obtain genomic characteristics of pathogens quickly and accurately. An easy-to-use and convenient web service platform, such as BacWGSTdb (<http://bacdb.org/BacWGSTdb>) [27], can provide uniform classification criteria to help us identify the taxonomy of species, ARGs, and virulence genes. Moreover, worldwide species information can be integrated into the platform, helping us to evaluate the tracking and evolution of pathogens.

Annotation of genes in the genome supports that AF18 possesses many common features to the *Enterobacteriaceae* family, such as the flagella that confer mobility to the bacterium, many genetic mobile elements that facilitate the transfer of stress-adaptive genes, especially ARGs. In fact, strains of this species have been identified as an important source of extended-spectrum β -lactamase-encoding genes and even carbapenem-resistant genes. For example, *Enterobacteriaceae* bacterium ENNIH2 is KPC-2 positive [24], *Metakosakonia* sp. MRY16–398 carries *bla*_{IMP-6}, *bla*_{CTX-M-2}, and *aadA2* gene [23], and the [*Kluyvera*] *intestini* str. GT-16 contains many ARGs that even confer resistance to polymyxin [25]. A more distant relative of AF18—*Pyrobacter ursingii* (previously named *Kluyvera intermedia*) was found to be KPC positive and carbapenem-resistant [28]. AF18 carries genes of the *hcp-clp* and *mprAB* system which are powerful in implementing persistence status and endows resistance to many environmental stresses including all kinds of antibiotics [29, 30]. These species can acquire various ARGs from their *Enterobacteriaceae* relatives in natural environments and host gut. For example, *E. coli* and *K. pneumoniae* can share these

ARGs through various genetic mobile elements or even in a more efficient manner of conjugating resistant plasmids [31]. Thus, the species of AF18 may prove important in spreading ARGs and function as a mediator in the eco-evolutionary feedback loops of AMR. In this regard, the novel species and many other emerging opportunistic pathogens of *Enterobacteriaceae* family, such as *Kluyvera* spp. and *Enterobacter* spp., deserve more attention in clinical practice and in the field of antibiotic resistance control.

In our study, the resistant AF18 does not have to transfer its resistance gene to the co-infected sensitive *K. pneumoniae* to confer protection. Being antibiotic-resistant by itself, we speculate that AF18 just upregulated the production of antibiotic-hydrolase to generate a niche of low antibiotic concentration for the sensitive *K. pneumoniae* to hide. Such co-operation between different bacteria has been deeply investigated in experiments, and relevant theories or mathematic models have been fully developed which illustrated the important role of growth regulation in maximizing the benefit and population of the entire consortium [32–34]. Our study provides empirical evidence for these hypotheses and highlights the importance of mutualistic relationships between microbes in a co-infection scenario in clinical settings. Transcriptome analysis further identified genes involved in growth regulation and pointed to two novel sRNAs that might be the key regulators of the process. As antibiotics are not always successful, especially in treating opportunistic pathogens, the sRNAs that promote the growth of host strains as we had identified may serve as targets of bacteriostatic agents and deserve further investigations.

Conclusion

Opportunistic pathogenic strain *Enterobacteriaceae bacterium* AF18 is in a novel species with little documentation. In-depth genomic analysis suggests the ability of this species to transfer between the natural environment and host habitats and in disseminating antibiotic-resistant genes, which potentiates it to be an important disseminator of antibiotic resistance to clinical pathogens. When co-infected with typical pathogens, the resistant opportunistic strain can provide temporary protection for the whole consortium and cause confusions in the etiological diagnosis and antibiotic treatment, although the strain by itself is possibly not a pernicious pathogen for immunocompetent patients. Taken together, *Enterobacteriaceae bacterium* AF18 and other newly emerging opportunistic pathogens complicate the situation of antibiotic resistance control in clinical practice and deserve in-depth investigation including methods for surveillance and control.

Methods

Bacterial isolation

The bile sample was collected from a patient with obstructive jaundice who suffered an infection 2 days after the percutaneous transhepatic cholangial drainage (PTCD) surgery admitted to Fujian Province Hospital (N26.08, E119.30) in 2014. The original colonies isolated from the bile sample on a blood agar plate were further incubated in Luria-Bertani (LB) broth overnight. The culture was diluted 10⁶-fold with LB broth and spread onto MacConkey agar and incubated at 37 °C for 24 h. Then, the morphology of the isolated colonies were observed. The *K. pneumoniae* colonies were inoculated into LB broth and incubated overnight. The bacterial cultures were supplemented with glycerol before freezing and stored in a -80°C freezer. AF18 was selected by culturing overnight in LB broth containing 50 µg/ml ceftriaxone: the aforementioned method was used to pick 10 drug-resistant clones from the MacConkey agar medium; then, the clones were re-inoculated in LB broth containing 50 µg/ml ceftriaxone and cultured overnight. The selection was repeated 5 times, and each culture was supplemented with glycerol before freezing and stored in a -80°C freezer.

Biological characterization of strain AF18

Gram staining

(1) A total of 10 µl of a bacterial culture that was cultured for 4 h in LB broth without antibiotics was dropped onto a sterile slide. (2) The bacteria were stained for 1 min with ammonium oxalate crystal violet and the dye was washed off with water. (3) A potassium iodide solution was dropped onto the bacteria and allowed to stain for 1 min and the excess dye was gently rinsed off with water. (4) Next, 95% ethanol was dropped onto the slide to destain for 30 s. (5) Safranin solution was dropped onto the slide to counterstain for 30 s, and the dye solution was rinsed off with water. When the slide was dry, the bacteria were observed under a microscope with an oil objective lens.

Biochemical identification: The biochemical identification of AF18 was performed with the VITEK-II automated bacterial identification system and the API20E *Enterobacter* biochemical identification system (Biomérieux, France). The obtained biochemical properties were compared with the biochemical properties of known species. The VITEK-MS mass spectrometer and the time-of-flight detection method were used to obtain the mass spectrum of AF18, which was compared with the mass spectra of known species. Antimicrobial susceptibility of strain AF18 and *K. pneumoniae* was investigated by broth microdilution using the E-test (Biomérieux, France) according to manufacturers' instructions, and a total of 21 antibiotics were tested as listed in Table 1.

Electron microscopy

An inoculation loop was used to scrape a bacterial colony from a blood agar plate. The colony was fixed with 2.5% glutaraldehyde, followed by 1% osmium acid, for 2 h each. After gradient dehydration with 30, 50, 70, 80, 90, and 100% ethanol, the bacteria were dried in a desiccator. After the specimen was coated with gold, the morphology of the bacteria was observed under a Hitachi S-3400 N scanning electron microscope. Bacteria grown on a blood agar plate for approximately 6 h were scraped with an inoculation loop. A bacterial suspension prepared in saline was dropped onto a copper net and allowed to rest for 5 min so that the bacteria could adhere to the copper net; then, the excess water was removed. The specimen was stained with 1% uranyl acetate for approximately 10 min. After the specimen dried, the flagella were observed with a JEM-1230 transmission electron microscope.

Genome sequencing

AF18 cells were harvested by centrifugation from an overnight culture, and DNA was extracted with QIAamp DNA Mini Kit (Qiagen, Cat No: 51304) following manufacturer's instruction. The extracted DNA sample was assayed with a NanoDrop spectrophotometer for quantification and then sent to Beijing Novogene Bioinformatics Technology Co., LTD for whole-genome sequencing. The genome of AF18 was both sequenced with the single-molecule real-time sequencing (SMRT) technology from Pacbio using a PacBio RSII sequencer (insert size was approximately 10 kb) for one cell, and sequenced with the short-reads HiSeq 2000 platform from Illumina (100 bp pair-end reads) for 3G raw reads. The obtained raw SMRT reads were analyzed and de novo assembled using SMRT Analysis 2.3.0 software. Then, we performed the error correction of tentative complete circular sequences using Pilon (v1.18) with Illumina short reads.

Genome annotation

GeneMarks (version 4.17) [35] was used to predict protein-coding genes. The Kyoto Encyclopedia of Genes and Genomes (KEGG), Clusters of Orthologous Groups (COG), NCBI-NR and Gene Ontology (GO) databases were used to annotate the function of predicted genes [36–38]. Virulent genes were reliably identified by BLAST in the Virulence Factors Database (<http://www.mgc.ac.cn/VFs/>) with identity > 80% and E-value < 10e-50, and antibiotic-resistant genes were identified using RGI from the Comprehensive Antibiotic Resistance Database (<https://card.mcmaster.ca/>) with “perfect and strict hits” and “identity > 90%” [39, 40]. Circos software (Version 0.64) was used to plot a circular map of the genome [41]. Plasmid

replicon typing was performed using the curated PlasmidFinder database at the CGE website (<https://cge.cbs.dtu.dk/services/PlasmidFinder/>) [42].

Phylogenetic analysis

We used Mash [16] to compare the AF18 genome sequence to the NCBI genome assembly database, and picked 33 genomes from the non-redundant species with identity scores > 75%. Then, we used kSNP3 to identify core genome SNPs (single nucleotide polymorphisms) of each pair of the 34 genome sequences with an optimal k-mer size of 21 (determined by Kchooser) [43]. These core SNPs were used to build a maximum likelihood tree by FastTreeML [44], and iTOL was used to exhibit the phylogenetic tree (<https://itol.embl.de/>) [45]. We used fastANI to calculate the pairwise ANI of the 34 genome sequences [17]. We used RNAmmer to predict the rRNA sequences [46], and then used BLASTN to align the 16S rRNA sequences to the genomes we picked.

Comparative analysis of pAF18_2

We used BLASTN to compare pAF18_2 sequence against the NR/NT database and picked 10 non-redundant plasmids according to the query coverage (> 60%) and percent identity (> 99.7%) [47]. We then performed BLASTN for pAF18_2 against the above selected 10 plasmids to find out the alignment fragments with E-value < e-50 [48], and generated the comparative map using CGView (<http://wishart.biology.ualberta.ca/cgview/>) [49].

Drug-resistant plasmid elimination test

A single colony of AF18 was inoculated into LB broth medium without antibiotics and cultured for 24 h at 37 °C. Then the culture was 1:1000 diluted and re-inoculated in LB broth medium for another 24 h. The procedure was repeated while an aliquot was collected and spread on LB agar medium for each round of re-inoculation. Colonies grown on the LB agar medium were randomly selected and tested for the presence of *bla*_{CTX-M-3} gene by PCR. DNA was extracted with QIAamp DNA Mini Kit, and PCR was performed under cycling conditions: 95 °C for 5 min, followed by 32 cycles of 95 °C for 1 min, 55 °C for 1 min, 72 °C for 1 min, followed by a single step of 72 °C for 5 min. The pair of Primers were: F 5'-CAGAATAAGGAATCCCATG-3', and R 5'-CGTCTAAGGCGATAAACA-3'. The PCR negative colonies, which might have lost their resistant plasmid pAF18-2, were functionally confirmed by inoculation in LB broth medium both with and without ceftriaxone (20 µg/ml). The strains that didn't survive in the ceftriaxone-containing medium were believed to have lost the resistant plasmid, and one of them was preserved and named AF18-NC.

Co-culture of AF18 and *K. pneumoniae*

AF18 and *K. pneumoniae* were first co-cultured in LB broth mediums containing different concentrations of ceftriaxone (0.125–16 µg/mL) overnight, and then the bacterial solution in the tube was diluted and spread on MacConkey agar medium. We picked out the colonies of *K. pneumoniae* from the plates and incubated them in LB broth medium containing ceftriaxone at the same concentration overnight, and then diluted the bacterial culture with LB broth and spread onto MacConkey agar medium.

Growth rate measurement

Overnight cultures of the *K. pneumoniae* strain, AF18, and AF18-NC were sampled and diluted to O.D. = 0.10, and were then cultured at 37 °C for 24 h in LB broth medium. The turbidity (O.D. value) of the cultures was measured by using a bacterial turbidimeter at 1 h intervals for 24 h, and the 3 bacterial growth curves were drawn based on the O.D. values. The proportion of AF18 and the *K. pneumoniae* strain in co-cultured samples was determined by colony counting on plates of MacConkey agar.

Transcriptome sequencing

AF18 and AF18-NC cells were harvested from its overnight LB broth medium culture by centrifugation, and total RNA of both strains were extracted with QIAGEN RNeasy Plus Mini Kit (Qiagen, Cat No.74134) following manufacturer's instruction. The extracted RNA samples were assayed with NanoDrop spectrophotometer for quantification and then sent to Beijing Novogene Co., LTD for transcriptome sequencing. rRNA was removed using the Ribo-Zero rRNA Removal Kit (Epicentre Biotechnologies), then the transcriptome was sequenced on Hiseq 2000 platform with 1G raw reads (paired-end, 2 × 100 bp). Low-quality reads and adaptor sequences were then removed. Using the whole genome of AF18 as the reference genome, the gene expression level for each transcript was estimated by calculating the FPKM (Fragments Per Kilobase per Million mapped fragments) value of each transcript.

Differential gene expression analysis

The read count data of each transcript was first normalized using DEseq [50]. According to the binomial distribution model, hypothesis testing was performed on each transcript between the AF18 strain and the AF18-NC strain, and confirmed by multiple hypothesis tests.

Functions of differentially expressed genes

Functions of the differentially expressed genes between AF18 and AF18-NC were annotated with GO database [36], and the probability of enrichment for each cluster

was calculated by using the weight algorithm and Fisher's exact test implemented in topGO package [51]. Clusters with a corrected *p* value < 0.05 were regarded as significantly enriched.

sRNA analysis

Rockhopper software was used to search new intergenic transcripts [52, 53], and those transcripts that had no hits in the NCBI-NR protein database by BLASTx were considered as candidates for non-coding sRNA. The sRNAs with two-fold increased/decreased FPKM value were regarded as up-/down-regulated. Secondary structures of candidate sRNAs were predicted using RNAfold software and their target genes were predicted by using IntaRNA [54, 55].

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-020-6720-z>.

Additional file 1: Figure S1. The bar plot of enriched GOs in AF18 and AF18-NC

Additional file 2: Figure S2. The sequences and the secondary structures of sRNAs

Additional file 3: Table S1. The results of biochemical testing of AF18 isolate by VITEK II2

Additional file 4: Table S2. Accession numbers of genome and plasmid sequences used for comparative analyses

Additional file 5: Table S3. Comparison of whole-genome ANI and 16S rDNA BLASTN results

Additional file 6: Table S4. The ARGs and virulence genes annotation results of AF18

Additional file 7: Table S5. The summary statistics of genomic features of AF18

Abbreviations

AMR: Antimicrobial resistance; ARGs: Antibiotic resistance genes; ANI: Average nucleotide identity; COG: Clusters of Orthologous Groups; FPKM: Fragments Per Kilobase per Million mapped fragments; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; LB: Luria-Bertani; NCBI: National Center for Biotechnology Information; PTCD: Percutaneous transhepatic cholangial drainage; SMRT: Single-molecule real-time sequencing; SNPs: Single nucleotide polymorphisms

Acknowledgements

We are thankful to Donghui Yan and Wei Sun from Department of Clinical Laboratory Center, Beijing Friendship Hospital, for their help on the biological characterization. We are also thankful to the staff of Electron Microscope Laboratory of China Agricultural University, for their help on morphological characterization.

Authors' contributions

ZG and YK designed and conducted the study. YC isolated the strains and organized the clinical information. ZZ, DL, YZ, YG, LZ, and YL Z performed the experiments. ZW and JS identified the isolate. ZZ, XS, and YH performed the bioinformatics analyses. ZZ, YK and XS prepared the manuscript. ZG revised the manuscript, and all authors read the manuscript and approved the submission.

Funding

The study was funded by National Science and Technology Major Project (2017ZX10103004-006, 2018ZX10712001-018-002), National Natural Science

Foundation of China (81870010, 31970568, 31671350). Programs of the Chinese Academy of Sciences (Y8YZ02E001, QYZDY-SSW-SMC017). The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Availability of data and materials

The complete, annotated genomic sequence of AF18 was deposited in a public database GenBank [56] (accession numbers: chromosome, CP025982; pAF18_1, CP025983; pAF18_2, CP025984). The clean sequence data for RNA-Seq have been deposited in the Genome Sequence Archive [57] in BIG Data Center [58], Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under accession number CRA002037 that are publicly accessible at <https://bigd.big.ac.cn/gsa>. Other genome and plasmid sequences used for comparative analyses in this study are publicly available in the NCBI Assembly database (<https://www.ncbi.nlm.nih.gov/assembly>) and Nucleotide database (<https://www.ncbi.nlm.nih.gov/nucleotide>) under accession numbers listed in Table S2.

Ethics approval and consent to participate

The study protocol was approved by the ethics committee of Peking University People's Hospital (Approval No. 2015PHB037-01, 17/01/2015). Written consent was acquired from the patient.

Consent for publication

Not applicable.

Competing interests

All authors declare that they have no conflicts of interest.

Author details

¹Department of Respiratory & Critical Care Medicine, Peking University People's Hospital, Beijing, China. ²Department of Respiratory & Critical Care Medicine, the First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China. ³Department of Respiratory and Critical Care Medicine, Zhongshan Hospital Xiamen University, Xiamen 361004, Fujian, China. ⁴Beijing Institute of Genomics, Chinese Academy of Sciences, Beijing, Beijing, China. ⁵University of Technology Sydney, Ultimo, NSW, Australia. ⁶Department of Respiratory, Critical Care and Sleep Medicine, Xiamen University Xiang'an Hospital, Xiamen, Fujian, China. ⁷Department of Respiratory & Critical Care Medicine, Fujian Provincial Hospital, Fuzhou, Fujian, China. ⁸Laboratory Medicine, Peking University People's Hospital, Beijing, China. ⁹Department of Clinical Laboratory Center, Beijing Friendship Hospital, Beijing, Beijing, China.

Received: 20 January 2020 Accepted: 5 April 2020

Published online: 15 April 2020

References

- O'Neill J. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. *Rev Antimicrob Resist*. 2014;20:1–16.
- Woolhouse M, Ward M, van Bunnik B, Farrar J. Antimicrobial resistance in humans, livestock and the wider environment. *Philos Trans R Soc Lond Ser B Biol Sci*. 2015;370(1670):20140083.
- Mourkas E, Florez-Cuadrado D, Pascoe B, Calland JK, Bayliss SC, Mageiros L, Meric G, Hitchings MD, Quesada A, Porrero C, et al. Gene pool transmission of multidrug resistance among campylobacter from livestock, sewage and human disease. *Environ Microbiol*. 2019;21(12):4597–613.
- Hassell JM, Ward MJ, Muloi D, Bettridge JM, Phan H, Robinson TP, Ogendo A, Imboma T, Kiiru J, Kariuki S, et al. Deterministic processes structure bacterial genetic communities across an urban landscape. *Nat Commun*. 2019;10(1):2643.
- Hiltunen T, Virta M, Laine AL. Antibiotic resistance in the wild: an eco-evolutionary perspective. *Philos Trans R Soc Lond Ser B Biol Sci*. 2017;372(1712):20160039.
- Woolhouse ME, Ward MJ. Microbiology. Sources of antimicrobial resistance. *Science*. 2013;341(6153):1460–1.
- Ekundayo TC, Okoh AI. Pathogenomics of virulence traits of *Plesiomonas shigelloides* that were deemed inconclusive by traditional experimental approaches. *Front Microbiol*. 2018;9:3077.
- Romano-Bertrand S, Bourdier A, Aujoulat F, Michon AL, Masnou A, Parer S, Marchandin H, Jumas-Bilak E. Skin microbiota is the main reservoir of *Roseomonas mucosa*, an emerging opportunistic pathogen so far assumed to be environmental. *Clin Microbiol Infect*. 2016;22(8):737.e731–7.
- Schurch AC, van Schaik W. Challenges and opportunities for whole-genome sequencing-based surveillance of antibiotic resistance. *Ann N Y Acad Sci*. 2017;1388(1):108–20.
- Hunter CJ, Bean JF. Cronobacter: an emerging opportunistic pathogen associated with neonatal meningitis, sepsis and necrotizing enterocolitis. *J Perinatol*. 2013;33(8):581–5.
- Shin H, Lee JH, Choi Y, Ryu S. Complete genome sequence of the opportunistic food-borne pathogen *Cronobacter sakazakii* ES15. *J Bacteriol*. 2012;194(16):4438–9.
- Taylor LH, Latham SM, Woolhouse ME. Risk factors for human disease emergence. *Philos Trans R Soc Lond Ser B Biol Sci*. 2001;356(1411):983–9.
- Mather AE, Reid SW, Maskell DJ, Parkhill J, Fookes MC, Harris SR, Brown DJ, Coia JE, Mulvey MR, Gilmour MW, et al. Distinguishable epidemics of multidrug-resistant *Salmonella* Typhimurium DT104 in different hosts. *Science*. 2013;341(6153):1514–7.
- Iredell J, Brown J, Tagg K. Antibiotic resistance in Enterobacteriaceae: mechanisms and clinical implications. *Bmj*. 2016;352:h6420.
- Adeolu M, Alnajjar S, Naushad S. R SG: genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacteriales Ord. Nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. Nov., Pectobacteriaceae fam. Nov., Yersiniaceae fam. Nov., Hafniaceae fam. Nov., Morganellaceae fam. Nov., and Budviciaceae fam. Nov. *Int J Syst Evol Microbiol*. 2016;66(12):5575–99.
- Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. 2016;17(1):132.
- Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9(1):5114.
- Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun*. 2019;10(1):5029.
- Bouma JE, Lenski RE. Evolution of a bacteria/plasmid association. *Nature*. 1988;335(6188):351–2.
- Bengoechea JA, Sa Pessoa J. *Klebsiella pneumoniae* infection biology: living to counteract host defences. *FEMS Microbiol Rev*. 2019;43(2):123–44.
- Alnajjar S, Gupta RS. Phylogenomics and comparative genomic studies delineate six main clades within the family Enterobacteriaceae and support the reclassification of several polyphyletic members of the family. *Infect Genet Evol*. 2017;54:108–27.
- Tetz G, Tetz V. Draft Genome Sequence of *Kluyvera intestini* Strain GT-16 Isolated from the Stomach of a Patient with Gastric Cancer. *Genome Announc*. 2016;4(6).
- Sekizuka T, Matsui M, Takahashi T, Hayashi M, Suzuki S, Tokaji A, Kuroda M. Complete Genome Sequence of bla_{IMP-6}-Positive *Metakosakonia* sp. MR16–398 Isolate From the Ascites of a Diverticulitis Patient. *Front Microbiol*. 2018;9:2853.
- Weingarten RA, Johnson RC, Conlan S, Ramsburg AM, Dekker JP, Lau AF, Khil P, Odom RT, Deming C, Park M, et al. Genomic Analysis of Hospital Plumbing Reveals Diverse Reservoir of Bacterial Plasmids Conferring Carbapenem Resistance. *MBio*. 2018;9(1).
- Tetz G, Vechevskovskaya M, Zappile P, Dolgalev I, Tsigirig A, Heguy A, Tetz V. Complete Genome Sequence of *Kluyvera intestini* sp. nov., Isolated from the Stomach of a Patient with Gastric Cancer. *Genome Announc*. 2017;5(43).
- Benslama O, Boulahrouf A. High-quality draft genome sequence of *Enterobacter* sp. Bisph2, a glyphosate-degrading bacterium isolated from a sandy soil of Biskra, Algeria. *Genom Data*. 2016;8:61–6.
- Ruan Z, Feng Y. BacWGSTdb, a database for genotyping and source tracking bacterial pathogens. *Nucleic Acids Res*. 2016;44(D1):D682–7.
- Sheppard AE, Stoesser N, Wilson DJ, Sebra R, Kasarskian A, Anson LW, Giess A, Pankhurst LJ, Vaughan A, Grim CJ, et al. Nested Russian doll-like genetic mobility drives rapid dissemination of the Carbapenem resistance gene bla_{KPC}. *Antimicrob Agents Chemother*. 2016;60(6):3767–78.
- He H, Hovey R, Kane J, Singh V, Zahrt TC. MprAB is a stress-responsive two-component system that directly regulates expression of sigma factors SigB and SigE in mycobacterium tuberculosis. *J Bacteriol*. 2006;188(6):2134–43.
- Wu HY, Chung PC, Shih HW, Wen SR, Lai EM. Secretome analysis uncovers an Hcp-family protein secreted via a type VI secretion system in *agrobacterium tumefaciens*. *J Bacteriol*. 2008;190(8):2841–50.

31. Gazouli M, Tzouveleki LS, Prinarakis E, Miriagou V, Tzelepi E. Transferable cefoxitin resistance in enterobacteria from Greek hospitals and characterization of a plasmid-mediated group 1 beta-lactamase (LAT-2). *Antimicrob Agents Chemother.* 1996;40(7):1736–40.
32. Scherlach K, Hertweck C. Mediators of mutualistic microbe-microbe interactions. *Nat Prod Rep.* 2018;35(4):303–8.
33. Li L, Wu T, Wang Y, Ran M, Kang Y, Ouyang Q, Luo C. Spatial coordination in a mutually beneficial bacterial community enhances its antibiotic resistance. *Commun Biol.* 2019;2:301.
34. Wang X, Kang Y, Luo C, Zhao T, Liu L, Jiang X, Fu R, An S, Chen J, Jiang N, et al. Heteroresistance at the single-cell level: adapting to antibiotic stress through a population-based strategy and growth-controlled interphenotypic coordination. *MBio.* 2014;5(1):e00942–13.
35. Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 2001;29(12):2607–18.
36. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, et al. The gene ontology (GO) database and informatics resource. *Nucleic Acids Res.* 2004;32(Database issue):D258–61.
37. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28(1):33–6.
38. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
39. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, Lago BA, Dave BM, Pereira S, Sharma AN, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 2017;45(D1):D566–73.
40. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 2019;47(D1):D687–92.
41. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. Circos: an information aesthetic for comparative genomics. *Genome Res.* 2009;19(9):1639–45.
42. Carattoli A, Hasman H. PlasmidFinder and in Silico pMLST: identification and typing of plasmid replicons in whole-genome sequencing (WGS). *Methods Mol Biol.* 2020;2075:285–94.
43. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics.* 2015;31(17):2877–8.
44. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One.* 2010;5(3):e9490.
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.* 2016;44(W1):W242–5.
46. Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 2007;35(9):3100–8.
47. Johnson M, Zaretskaya I, Raytselis Y, Merezukh Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 2008;36(Web Server issue):W5–9.
48. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421.
49. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. *Bioinformatics.* 2005;21(4):537–9.
50. Anders S, Huber W. Differential expression of RNA-Seq data at the gene level—the DESeq package. Heidelberg: European Molecular Biology Laboratory (EMBL); 2012.
51. Alexa A, Rahnenfuhrer J. Gene set enrichment analysis with topGO. *Bioconductor Improv.* 27; 2019.
52. McClure R, Balasubramanian D, Sun Y, Bobrovskyy M, Sumbly P, Genco CA, Vanderpool CK, Tjaden B. Computational analysis of bacterial RNA-Seq data. *Nucleic Acids Res.* 2013;41(14):e140.
53. Tjaden B. De novo assembly of bacterial transcriptomes from RNA-seq data. *Genome Biol.* 2015;16:1.
54. Hofacker IL, Fontana W, Stadler PF, Bonhoeffer LS, Tacker M, Schuster P. Fast folding and comparison of RNA secondary structures. *Monatshefte für Chemie/Chemical Monthly.* 1994;125(2):167–88.
55. Mann M, Wright PR, Backofen R. IntaRNA 2.0: enhanced and customizable prediction of RNA-RNA interactions. *Nucleic Acids Res.* 2017;45(W1):W435–9.
56. Sayers EW, Cavanaugh M, Clark K, Ostell J, Pruitt KD, Karsch-Mizrachi I. GenBank. *Nucleic Acids Res.* 2019;47(D1):D94–9.
57. Wang Y, Song F, Zhu J, Zhang S, Yang Y, Chen T, Tang B, Dong L, Ding N, Zhang Q, et al. GSA: genome sequence archive. *Genom Proteomics Bioinform.* 2017;15(1):14–8.
58. National Genomics Data Center M. Partners: database resources of the National Genomics Data Center in 2020. *Nucleic Acids Res.* 2020;48(D1):D24–33.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

