



## **An insight into the genome of extensively drug-resistant and uropathogenic *Citrobacter werkmanii***

Shadab Parvez, Asad U Khan, Gurwinder Kaur, mohamed Barakat, Philippe Ortet, Shanmugam Mayilraj

### **► To cite this version:**

Shadab Parvez, Asad U Khan, Gurwinder Kaur, mohamed Barakat, Philippe Ortet, et al.. An insight into the genome of extensively drug-resistant and uropathogenic *Citrobacter werkmanii*. *Journal of Global Antimicrobial Resistance*, 2020, 22, pp.785 - 791. <10.1016/j.jgar.2020.06.014>. <hal-03071675>

**HAL Id: hal-03071675**

**<https://hal.science/hal-03071675v1>**

Submitted on 16 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Journal Pre-proof

An insight into the genome of extensively drug-resistant and uropathogenic *Citrobacter werkmanii*

Shadab Parvez, Asad U. Khan, Gurwinder Kaur, Mohamed Barakat, Philippe Ortet, Shanmugam Mayilraj



PII: S2213-7165(20)30156-9  
DOI: <https://doi.org/10.1016/j.jgar.2020.06.014>  
Reference: JGAR 1282  
To appear in: *Journal of Global Antimicrobial Resistance*  
Received Date: 7 April 2020  
Revised Date: 28 May 2020  
Accepted Date: 18 June 2020

Please cite this article as: Parvez S, Khan AU, Kaur G, Barakat M, Ortet P, Mayilraj S, An insight into the genome of extensively drug-resistant and uropathogenic *Citrobacter werkmanii*, *Journal of Global Antimicrobial Resistance* (2020), doi: <https://doi.org/10.1016/j.jgar.2020.06.014>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Published by Elsevier.

**An insight into the genome of extensively drug-resistant and uropathogenic *Citrobacter werkmanii***

**Running title:** Genome analysis of drug-resistant *C. werkmanii*.

Shadab Parvez<sup>1,2</sup>, Asad U. Khan<sup>1\*</sup>, Gurwinder Kaur<sup>3</sup>, Mohamed Barakat<sup>4</sup>, Philippe Ortet<sup>4</sup>, Shanmugam Mayilraj<sup>3</sup>

<sup>1</sup>Medical Microbiology and Molecular Biology Lab., Interdisciplinary Biotechnology Unit, A.M.U., Aligarh – 202 002, India. <sup>2</sup>CIF, S. Ramanujan Block, Mujeeb Bagh, JMI, New Delhi – 110 025. <sup>3</sup>MTCC- Microbial Type Culture Collection & Gene Bank, CSIR- Institute of Microbial Technology, Chandigarh - 160 036, India. <sup>4</sup>Aix-Marseille University, CEA, CNRS, LEMiRE, UMR 7265, BIAM, Saint-Paul-lez-Durance, France.

**\*Corresponding author:** Prof.Asad U. Khan, Medical Microbiology and Molecular Biology Lab., Interdisciplinary Biotechnology Unit, Aligarh Muslim University, Aligarh – 202 002, UP, India. Ph.: 0091-9837021912, Fax: 0091-571-2721776

Email id: asad.k@rediffmail.com

**Highlights**

- First report on whole genome analysis of extensively drug-resistant NDM-6 producing uropathogenic *C. werkmanii*, ST-104.
- Genome analysis of AK-8 led us to identify BaeSR two-component system, regulating the production of multidrug efflux proteins.
- ntBLAST results showed AK-8 to possess 69.60% genetic identity with *E. coli*536
- AK-8 genome encodes 21 chemoreceptors that help in colonization and pathogenesis

## Abstract

**Objective:** Carbapenemase producing bacteria poses serious threat to public. This study was performed to understand the emergence and genetic features of NDM-producers in hospital setting.

**Materials and Methods:** Samples were collected from tertiary care hospital. Isolate identification was carried out by 16S rRNA sequencing. The genome of *Citrobacter werkmanii* (AK-8) was sequenced on Illumina-NextSeq-500 platform. Resistant determinants and pathogenicity islands were determined by ResFinder and PathogenFinder, respectively. MLST, two-component system and transcription factors were identified by P2RP server, while, variant calling and insertion sequences were determined by Galaxy and ISfinder, respectively. AK-8 genome was compared with *Escherichia coli* (536-UPEC).

**Results:** It is the first report on whole genome analysis of extensively drug-resistant NDM-6 producing uropathogenic *Citrobacter werkmanii*, ST-104. Resistance genes for all antibiotics except colistin, fosfomycin, fusidic-acid, nitroimidazole, oxazolidinones, tetracycline and glycopeptides were detected in the isolated strain. The genome analysis of AK-8 led us to identify BaeSR two-component system, regulating the production of multidrug efflux proteins. The virulence was found to be regulated by CpxRA, ZraRS, RstAB, UhpAB, AcrAB, RcsBc, and UvrY whereas, Bar-UvrY was found to control carbon metabolism, flagellum biosynthesis and biofilm formation. AK-8 genome encodes 21 chemoreceptors, helping in colonization and pathogenesis. Fur family transcriptional regulator, cAMP receptor protein and RpoS were found to increase virulence of AK-8. ntBLAST analysis showed 69.60% genetic identity with *E. coli* 536 as an adaptive feature acquired for survival.

**Conclusion:** The emergence of extensively drug-resistant and pathogenic *Citrobacter werkmanii* is alarming and it should not be ignored as commensal.

**Keywords:** NDM-6, *Citrobacter werkmanii*, IS30, two-component system, uropathogen, antimicrobial resistance.

## 1. Introduction

Antibiotics are the most important discovery in medicine decreasing morbidity and mortality. These are crucial public health tools in acute care hospitals and long-term care facilities; providing treatment from infections such as central line-associated bloodstream infections, catheter-associated urinary tract infections, surgical site infections, etc., saving millions of lives worldwide. Today, the emergence of antibiotic resistance in bacteria has threatened the efficacy of antibiotics. Thereby, limiting the treatment options, the available drug choices are expensive and, in some cases, non-existing. Thus, posing one of the greatest threats to the healthcare setting across the globe (1).

Recently, a field report from the Centres for Disease Control and Prevention (CDC) documented a story of a 70-year-old female of Washoe County with travelling and hospitalization history in India. According to the report, in August 2016, Nevada public health officials primarily diagnosed her with systemic inflammatory response syndrome due to her femur fracture. Later, in September 2016, it caused a septic shock followed by her death. The isolated sample analysis at CDC confirmed that she was infected with a New Delhi metallo- $\beta$ -lactamase (NDM) producing *Klebsiella pneumoniae* (2).

The NDM enzyme was first identified in 2008 from a Swedish patient when he returned from New Delhi, India (3). Till date 28 variants of this enzyme have been reported (<http://www.lahey.org/Studies/other.asp#table1>) (4). The NDM variants producing *Enterobacteriaceae* are pan-resistant to  $\beta$ -lactams including carbapenems, which are often considered as antibiotic of last resort for treatment (5). Consequently, the curable infectious diseases have become hard-to-treat or untreatable. NDM-producing

*Enterobacteriaceae* have been identified in more than 70 countries so far (6). This has brought humanity at the “Dawn of a Post-Antibiotic Era” which poses a threat to public health across the globe.

Chronic kidney disease (CKD) has recently been recognized as a public health priority, globally (7,8). In 1990, based on the total number of deaths worldwide, CKD was ranked 27th (annual death rate 15.7 per 100,000) and in 2010 the position climbed to 18th (annual death rate 16.3 per 100,000) (7). This substantially burdens the economy of the country. In 2007, the US Medicare expenditure on CKD patients was more than US\$70 billion per year (8) whereas, UK National Health Service had estimated about £1.44 – 1.45 billion per year in 2009 – 10 (9).

The *Citrobacter werkmanii* is a member of *Enterobacteriaceae*, commonly found as commensal in humans and animals. The *Citrobacter* species are rare opportunistic bacteria found to be responsible for urinary tract infections, respiratory tract infections etc., in neonates and immunocompromised persons (10). In the present study, we analysed the genome of NDM-6 producing *Citrobacter werkmanii* isolated from a CKD patient aiming at establishing it as an extensively drug-resistant and uropathogen.

## **2. Materials and Methods**

### *2.1. Collection of clinical samples and case report*

In this study, 60 isolates were collected (Aug. 2013 to Sept. 2013) from patients admitted to the ICU, JN Medical College, and Hospital, Aligarh, India. One isolate from each patient was collected either from urine, pus or stool. The preliminary study showed that two isolates were carbapenem-resistant. However, only one isolate was stable carbapenemase producer whereas, rest of the 58 isolates were found negative for carbapenemase production in that time window. The stable carbapenemase

producer was isolated from urine of a 57-year-old patient, a male farmer who was a resident of Aligarh, India and was admitted to emergency department of the hospital and later shifted to ICU of medicine with the complaints of high grade fever, decreased urine output, shortness of breath and respiratory distress with decreased level of consciousness on the onset of admission. He was also complaining of burning sensation during urination. He was a known case of type 2 diabetes mellitus (2DM) and hypertension for the last 12 years and was on medication. Two years back, this patient was also diagnosed with chronic kidney disease (CKD). A urine sample was checked for infection and was found to have severe urinary tract infection (UTI). The final diagnosis was made as for type 2DM with systemic hypertension, acute on CKD and UTI.

Ceftriaxone was started as a conservative treatment for acute CKD. No improvement was found until next day and hence hemodialysis was initiated. The antibiotic was changed to cefoperazone and urine was sent for culture. Even after 2 days, there was no improvement, fever was persistent and renal function was further deteriorated. Linezolid was added to the treatment. Despite aggressive treatment and hemodialysis, the patient expired on the 6th day of admission. Urine culture report revealed that patient was infected with carbapenem-resistant *Enterobacteriaceae* (named as AK-8).

## 2.2. Isolate identification

The 16S rRNA sequence was used to predict isolate (AK-8) at the species level.

## 2.3. Whole genome sequencing (WGS)

Whole genome sequencing was performed on the total DNA isolated from AK-8 strain using Qiagen's QIAamp DNA mini kit (Cat# 51304). The purity and concentration of extracted DNA were determined using the GE SimpliNano UV-Vis Spectrophotometer. For sequencing, the DNA was run on Illumina NextSeq 500

platform for paired-end 150-bp sequencing. The de novo assembly of clean high-quality reads was performed by CLC Genomics Workbench ver. 9.0 (Qiagen, Germany). The BLAST of longest contig 1,065,488 bp at National Centre for Biotechnology Information (NCBI) got hit with *Citrobacter werkmanii* with 99.66% identity. Thus, this strain was used as a reference strain to generate reference assembly for the AK-8. The genome annotation was done by the Rapid Annotations using Subsystems Technology (RAST) server (11) and tRNAscan-SE-1.23 software (12).

To locate the *bla*<sub>NDM-6</sub> in the genome of AK-8, a stand-alone BLAST was performed against the assembled genome of AK-8 using nucleotide sequence of NDM-6 as a query. The potential ORFs (open reading frames) were predicted using the NCBI ORF finder program with minimum ORF length of 300 bases (<https://www.ncbi.nlm.nih.gov/orffinder/>).

#### 2.4. Identification of resistance determinants

Antibiotic resistance genes present in *Citrobacter werkmanii* was determined from WGS data using ResFinder web server, version 2.1 at the threshold of 98.00% identity (ID) ([www.genomicepidemiology.org](http://www.genomicepidemiology.org)) (13).

#### 2.5. Prediction of pathogenicity

The PathogenFinder web server version 1.1 was used to predict the potential pathogenicity of *Citrobacter werkmanii* (AK-8) ([www.genomicepidemiology.org](http://www.genomicepidemiology.org)) (14).

#### 2.6. Multilocus Sequence Typing (MLST) of *Citrobacter werkmanii* (AK-8)

MLST profile of *Citrobacter werkmanii* (AK-8) was determined from WGS data i.e. Illumina paired-end fastq files, using MLST web server, version 1.8 ([www.genomicepidemiology.org](http://www.genomicepidemiology.org)) (15).



### 2.7. Analysis of Two-Component System (TCS) and transcription factors (TFs)

Two-component system (TCS) and transcription factors (TFs) were predicted using the web server P2RP (<http://www.p2rp.org>) (16), which computes regulatory proteins (RPs) by taking as input genome DNA contigs. The sequences are first run through MED-Start (17), a prokaryotic gene-finding algorithm, to define a set of potential gene sequences. The putative genes are then translated to constitute a proteome for regulatory protein (RP) prediction. The identification of RP candidates in protein sequences was accomplished by domain analysis of each predicted protein, using RPSBLAST, as previously developed for P2CS and P2TF(18,19). The pool of domains used to search for RP proteins was manually selected from the literature and extracted from within Pfam (20) and SMART (21) databases. P2RP is a set of PHP scripts, designed to search the numerous combinations of RP modules and to categorize RP proteins into families based on similarity and/or domain architecture. Finally, the cellular localization of each TCS protein was determined by the presence or absence of transmembrane (TM) segments, using the HMMTOP predictor (22).

HKs (Histidine Kinases) and their cognate RRs (Response Regulators) are often encoded on adjacent genes on the DNA. Therefore, all gene clusters containing at least one HK and one RR gene were considered to encode functional HK-RR gene pairs and were thus considered to encode a specific TCS. Single HK and RR genes were categorized as ‘orphans’.

### 2.8. Genomic heterogeneity of *Citrobacter werkmanii* (AK-8)

Variant calling was performed by Galaxy: a web-based platform for the analysis of next-generation sequence data (<https://usegalaxy.org/>) (23)

### 2.9. Identification of insertion sequences

Insertion sequence elements were predicted by program ISfinder (<https://www-is.biotoul.fr/index.php>) (24).

### 2.10. Comparative genome analysis

The genome of *Citrobacter werkmanii* (AK-8) was compared with the genome of uropathogenic *Escherichia coli* strain 536 (UPEC) which is a model of extra-intestinal pathogenic *Escherichia coli* as per annotation by RAST server (25,11).

### 3. Results

#### 3.1. Identification of carbapenemase-producing isolate

Of 60 isolates analyzed, only two carbapenem-resistant strains were identified. However, only one strain was found to be a stable carbapenemase producer (AK-8), isolated from a CKD patient. The nucleotide blast alignment of its 16S rRNA sequence had shown that this organism was *Citrobacter werkmanii* (AK-8).

#### 3.2. Draft genome sequence of *Citrobacter werkmanii* (AK-8)

The genome of *Citrobacter werkmanii* (AK-8) was sequenced using Illumina NextSeq 500 platform for paired-end 150-bp sequencing technology. A total of 14,798,078 reads were used for assembly of the genome with the help of CLC Genomics Workbench (Version 9.0). The estimated coverage of *Citrobacter werkmanii* genome was 100-fold. A total of 47 contigs of 6,45,222,009 bp in length were constructed, with an N<sub>50</sub> of 537,644 bp. The final draft genome sequence consists of 5,222,009 bp with an average G+C content of 51.9 %. The whole-genome annotation showed that the AK-8 contains 5,154 genes, 5077 CDS, 177 pseudo genes, and 77 RNAs.

#### 3.3. Detection of resistance markers by ResFinder

ResFinder had predicted acquired genes that confer resistance to  $\beta$ -lactam [*bla*<sub>NDM-6</sub>, *bla*<sub>CTX-M-15</sub>, *bla*<sub>OXA-1</sub>, and *bla*<sub>OXA-10</sub>], aminoglycoside [*armA*, *aac*(3)*IIa*, *aadA1*, *aac*(6')*Ibcr*, *aadA5*, and *aph*(3')*VIa*], fluoroquinolones [*QnrB12*], macrolide, lincosamide, and streptogramin B [*msr*(*E*), *mph*(*A*), *mph*(*E*)], chloramphenicol

[*cmlA1*, *catB3*], rifampicin [*ARR3*], sulphonamide [*sul1*] and trimethoprim [*dfrA17*] in *Citrobacter werkmanii* genome (Fig. 1A and 1B).

Further, no gene was identified that could confer resistance in *Citrobacter werkmanii* against colistin, fosfomycin, fusidic acid, nitroimidazole, oxazolidinones, tetracycline, and glycopeptides (Fig. 1A and 1B).

### 3.4. Prediction of pathogenicity of *Citrobacter werkmanii* (AK-8)

The pathogenicity of *Citrobacter werkmanii* was predicted by comparing its proteins to a protein family database associated with pathogenic and non-pathogenic bacteria. The PathogenFinder predicted *Citrobacter werkmanii* to be a human pathogen (probability of being a human pathogen was 0.874). The proteins in AK-8 matched with 91 pathogenic families and 3 non-pathogenic families (Table S1).

### 3.5. Identification of MLST

MLST-1.8 web server showed 100% alignment with alleles e.g. arginine deiminase (*arcA\_37*), aspartate aminotransferase (*aspC\_58*), ATP-dependent Clp protease (*clpX\_59*), DNA primase (*dnaG\_14*), fatty acyl-CoA synthetase (*fadd\_65*), lysine-specific permease (*lysp\_53*), and malate dehydrogenase (*mdh\_14*) in the AK-8 genome. This predicted that *Citrobacter werkmanii* (AK-8) belonged to a sequence type, ST-104.

### 3.6. Identification of insertion sequences

The insertion sequences were annotated in the genome of *Citrobacter werkmanii* (AK-8) and it was found that AK-8 genome contains 116 IS elements (Table S2).

### 3.7. Analysis of the two-component systems (TCSs) and transcription factors (TFs)/ proteins in *Citrobacter werkmanii* (AK-8)

The *Citrobacter werkmanii* (AK-8) genome encodes 78 ORFs potentially involved in two-component signal transduction systems (TCSs) or His-Asp (Histidine-Aspartate) phosphorelay signalling. Of these 78 genes, 36 were predicted to encode histidine kinases (HKs) and 42 encode response regulators (RRs) (Table S3). Further, the AK-8 genome also encodes 366 DNA-binding proteins mostly belonging to Helix-Turn-Helix (HTH) domain [167 transcriptional regulators (TRs), 8 sigma factors (SFs), 19 other-DNA binding proteins (ODPs) and 138 one-component system (OCS) proteins] (Table S4).

### 3.8. Genomic diversity of *Citrobacter werkmanii* (AK-8)

The variant calling or variant identification was performed to detect genomic variation or differences between the sequence of uropathogenic *Citrobacter werkmanii* (AK-8) with reference to wild type *Citrobacter werkmanii* strain. We found 23 88 918 variations in the genome of AK-8 with respect to wild type *Citrobacter werkmanii*.

### 3.9. Comparative genome analysis of *Citrobacter werkmanii* (AK-8) with *Escherichia coli* strain 536

A total of 579 genetic entities were screened in the genome of *Citrobacter werkmanii* (AK-8) and the same was compared with uropathogenic *Escherichia coli* strain 536 (UPEC) genome which is a model of extraintestinal pathogenic *Escherichia coli* (25).

Genomic entities were studied under 5 categories:

1. Virulence, Disease and defence.
2. Phages, Prophages, Transposable elements, Plasmids.
3. Membrane Transport.

4. Iron acquisition and metabolism.

5. Stress Response.

These encompassed an array of genes for adhesion; bacteriocins, ribosomally synthesized antibacterial peptides; invasion and intracellular resistance; resistance to antibiotics and toxic compounds; phages, prophages; ABC transporters; cation transporters; protein and nucleoprotein secretion system, type IV; protein secretion system, type I; protein secretion system, type II; protein secretion system, type V; protein secretion system, type VI; protein secretion system, type VII (Chaperone/Usher pathway (CU)); protein secretion system, type VIII (Extracellular nucleation/precipitation pathway (ENP)); protein translocation across cytoplasmic membrane; TRAP transporters; uni-, sym-, and anti-porters; siderophores; cold shock; detoxification; heat shock; osmotic stress; oxidative stress; and periplasmic stress. (Table 1, Fig.2) (Table S5)

#### 4. Discussion

For the first time, we are documenting the whole genome analysis of extensively drug-resistant and uropathogenic NDM-6 producing *Citrobacter werkmanii* ST-104 (AK-8) strain isolated from a CKD patient. A study reported dual-risk ST101 lineage carrying KPC-2 and CTX-M-15 along with other antibiotic resistance determinants in hypervirulent *K. pneumoniae* unlike AK-8 which has ST-104 (26).

The genome analysis of *Citrobacter werkmanii* (AK-8) showed that this bacterium is pathogenic and multi-drug resistant. It was found to harbour genes that confer resistance to almost all clinically relevant drugs, including the first-line of empirical treatment drug of UTI such as trimethoprim-sulfamethoxazole and antibiotic of last resort for treatment i.e. carbapenems (27). Apparently, as per our genome analysis, the

drug that could have saved the life of this CKD patient was colistin, fosfomycin, fusidic acid, nitroimidazole, oxazolidinones, tetracycline and glycopeptides, as no genes were found that could confer resistance in AK-8 against them.

Furthermore, we found that *Citrobacter werkmanii* (AK-8) genome encodes 78 ORFs potentially involved in two-component signal transduction systems (TCSs) or His-Asp phosphorelation. Phosphorelation is a niche based adaptive response within host, classically described as the association of two proteins that communicate through a His-Asp. Briefly, an auto-phosphorylated histidine kinase (HK) sensor protein phosphorylates the receiver domain of a response regulator (RR) on a conserved Asp residue. In *Citrobacter werkmanii* (AK-8), among the 78 genes predicted to encode TCS proteins, 36 encode HKs and 42 encode RRs (Table S3). In AK-8, HKs encoding genome, which is related to cognate co-operonic RRs (Table S3) are mainly of the classic type (28). Relatively, in addition to the classical HK domain, a few TCS proteins are orphans, including 1 hybrid HK containing at least one REC domain. A kinase domain (2 unorthodox HKs) contains an additional histidine phosphotransferase domain (Hpt) and a NarL family response regulator (Table S3).

Moreover, the genome analysis led us to identify BaeSR two-component system (Table S3), known to regulate the production of multidrug efflux proteins in *Citrobacter werkmanii* (AK-8). This increases the multidrug and metal resistance by inducing the AcrD and MdtABC drug efflux systems (29).

In *Citrobacter werkmanii* (AK-8), we have also identified a cognate pair of TCS proteins CpxRA, as reported earlier in *Citrobacter rodentium* necessary for virulence (30). Our current analysis shows that in addition to CpxRA, there are five additional TCS pairs ZraRS, RstAB, UhpAB, ArcAB and RcsBC. Moreover, an orphan response regulator (UvrY) was also found in AK-8 as implicated previously in *Citrobacter*

*rodentium* virulence. The UvrY protein is the cognate response regulator for BarA in *E. coli* regulates the expression of *csrB* and *csrC* regulatory small RNAs (sRNAs). The *csrB* and *csrC* sRNAs bind to CsrA protein and prevent it from binding to the 5' untranslated region (UTR) of their target mRNAs which post-transcriptionally down- or upregulates the expression of target genes. CsrA controls carbon metabolism, flagellum biosynthesis, and biofilm formation (31-33).

Further analysis of the *Citrobacter werkmanii* (AK-8) genome revealed at least 21 genes encoding putative chemoreceptors, known as methyl-accepting chemotaxis proteins (MCPs) (Table S3), and two putative chemotaxis clusters designated as chemCI1 and chemCI2. Motility and the underlying chemotaxis system are important for the colonization and pathogenicity. In comparison, *C. koseri* and *C. Rodentium* carry a single chemotaxis system, whereas the *Citrobacter werkmanii* (AK-8) genome encodes a second chemotaxis cluster (chemCI2) that seems to be acquired via horizontal gene transfer. This is an important finding in the understanding of an additional pathway controlling traits other than motility (34) as supported by the identification of GGDEF and EAL domains in chemCI2. The GGDEF and EAL domains proteins are most abundant in prokaryotes and have sensory and signal transduction domains. This might be associated with regulation of virulence, bacterial morphogenesis and/or motility in AK-8 as reported earlier. The uropathogens like Uropathogenic *Escherichia coli* (UPEC) and in AK-8 (this study), causes infection by ascending to the urethra, colonizing in the bladder and further ascending the ureter and colonize in the kidney. This mechanism is expected to be modulated and monitored by GGDEF and EAL domains.

In addition to this, *Citrobacter werkmanii* (AK-8) genome encodes 366 DNA-binding proteins (Table S4) which mostly are Helix-Turn-Helix (HTH) domain-containing

proteins (167 transcriptional regulators, 8 sigma factors, 34 response regulators, 19 other-DNA binding proteins, and 138 one-component system proteins (OCSs)). These OCSs were defined as proteins carrying both input and output domains but are lacking phosphor-transfer domains, a characteristic feature of the two-component systems, as defined by Ulrich et al (35). A high number of DNA-binding proteins correspond to one-component transcription factors of the LysR family (36). It is also relevant to note that Gamma proteo-bacteria tend to have more TF families detected than any other bacteria with comparable numbers of TFs per genome (37). Further, genome analysis of AK-8 led us to identify a Fur family transcriptional regulator (Table S4), known to control virulence in pathogenic bacteria. The Fur's critical contribution to virulence may not be due to its classical role as a transcriptional repressor of metal acquisition but to its complex role as a transcriptional activator of virulence (38).

*Citrobacter werkmanii* (AK-8) carries three cAMP receptor protein (Crp)-coding genes. Crp is one of the most important transcriptional regulators, which can regulate large quantities of operons in different bacteria. It is the Crp and Crl's (Table S4) interaction with RpoS (Table S4) that regulates growth phase-specific modulation of RpoS, facilitating host colonization and virulence in numerous bacterial pathogens. RpoS (RNA polymerase, sigma S factor) offers to virulence through enhancing survival against host defense systems or directly regulating expression of virulence factors (39).

The Insertion sequences (ISs) are also one of the most important factors that help us to understand the genome organization and evolution (24). We have found a massive accumulation of IS elements in AK-8 genome. The Galaxy webs server and command line had shown that there are 2388918 variants in AK-8 genome (*Citrobacter*



*werkmanii* strain), implying its diversity. Acquiring such traits and variations over time could be an adaptive feature of *Citrobacter werkmanii* (AK-8).

We have compared the genomes of *Citrobacter werkmanii* (AK-8) and uropathogenic *Escherichia coli* strain 536 (UPEC), which is a model of extraintestinal pathogenic *Escherichia coli* in order to understand their relatedness (25). We have found an identity of 69.60% that may have been acquired through horizontal gene transfer mechanism. The genes for virulence, disease and defence in AK-8 share an identity of 72% with 17% uniqueness. Both strains were found to produce ribosomally synthesized colicin V. The colicins are proteinaceous toxin produced by members of *Enterobacteriaceae* e.g. *E. coli*, *Citrobacter freundii*, and *Shigella* (40). The *C. werkmanii* (AK-8) and *E. coli* 536 competitively produce this toxin for resource and space in urinary tract from other members of *Enterobacteriaceae*. In the urinary tract, sequestered or low concentrations of metals (as bacterial enzyme cofactor) limit its availability to microbes. This mechanism is called as nutritional immunity (41). Therefore, in order to cause infection, they have to circumvent the nutritional immunity of the host. In AK-8 and *E. coli* 536 genomes we have identified the subsystems that are involved in cobalt-zinc-cadmium resistance, mercury resistance operon, arsenic resistance, siderophores, and copper homeostasis (copper tolerance). These subsystems are predicted to overcome the nutritional immunity of the host. Furthermore, *C. werkmanii* (AK-8) harbour more genetic entities for horizontal gene transfer mechanism than *E. coli* 356 This allows AK-8 to acquire and disseminate genetic elements for survival in various ecological niches.

## 5. Conclusion

The emergence of *Citrobacter werkmanii* as a uropathogenic and extensively drug-resistant organism is alarming and it should not be ignored as commensal. The present

study gives us an insight into the resistance determinants, virulence factors, and their mode of communication in order to adapt to environment. Further research is required to identify the exact functions of all the genes by global transcriptional analysis, analysis of biochemical and catalytic properties of the expressed proteins, and phenotypic analysis of the mutants.

### **Competing interests**

The authors declare that they have no competing interests.

### **Ethics approval and consent to participate**

A formal consent from the institutional ethical committee was taken and clearance was obtained from the institute's ethics committee. Participants/guardians had provided written, informed consent to participate in the study. We have a specific format to get the consents of patients/parents of minors. These forms are confidential and cannot be disclosed as per Institutional ethics committee's guidelines. The name of committee/board is "Institutional Ethical Committee of Interdisciplinary Biotechnology Unit [Biot/307/01.06.13]", Aligarh Muslim University, Aligarh, India.

### **Availability of data and materials**

The sequence of NDM-6 was deposited in the NCBI database under the accession no. **KJ872581**. The whole genome NGS project was deposited at DDBJ/ENA/GenBank under the accession no. **NDXY00000000**.

### **Funding**

This work was supported by DBT (India) grants; BT/PR11453/BID/07/296/2009, and ICMR (India) grant AMR/5/2011-ECD-I to A.U.K. S.P.'s Ph.D. fellowship was awarded by UGC (BSR), Government of India.

## Acknowledgements

The authors are grateful to Dr. Rafiq and technical support at Medicine department, J N Medical College and Hospital, Aligarh, India, for providing samples and history of the patient. S.P. also acknowledge DST – PURSE program, CIF, S. Ramanujan Block, Jamia Millia Islamia, New Delhi – 25, India, for the award of post-doctoral fellowship. Ms Samantha Macchi, University of Arkansas at Little Rock, US, is acknowledged for correcting English.

## References

1. Schäberle, T.F., Hack, I.M. 2014. Overcoming the current deadlock in antibiotic research. *Trends Microbiol.* 22:165-7.
2. Chen, L., Todd, R., Kiehlbauch, J., et al. 2017. Notes from the Field : Pan-Resistant New Delhi Metallo-Beta-Lactamase-Producing *Klebsiella pneumoniae* — Washoe County, Nevada, 2016. *MMWR Morb. Mortal. Wkly. Rep.* 66:33.
3. Yong, D., Toleman, M.A., Giske, C.G., et al. 2009. Characterization of a New Metallo- $\beta$ -Lactamase Gene, blaNDM-1, and a Novel Erythromycin Esterase Gene Carried on a Unique Genetic Structure in *Klebsiella pneumoniae* Sequence Type 14 from India. *Antimicrob. Agents Chemother.* 53:5046–5054.
4. Liu, Z., Wang, Y., Walsh, T.R., et al. 2017. Plasmid-Mediated Novel blaNDM-17 Gene Encoding a Carbapenemase with Enhanced Activity in a ST48 *Escherichia coli* Strain. *Antimicrob. Agents Chemother.* 61:e02233-16.
5. Kumarasamy, K.K., Toleman, M.A., Walsh, T.R., et al. 2010. 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.
6. Woodford, N., Johnson, A.P. 2013. Global spread of antibiotic resistance: the example of New Delhi metallo- $\beta$ -lactamase (NDM)-mediated carbapenem resistance. *J. Med. Microbiol.* 62:499–513.
7. Lozano, R., Naghavi, M., Foreman, K., et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 380:2095–2128.
8. Jha, V., Garcia-Garcia, G., Iseki, K., et al. 2013. Chronic kidney disease: global dimension and perspectives. *Lancet* 382:260–272.

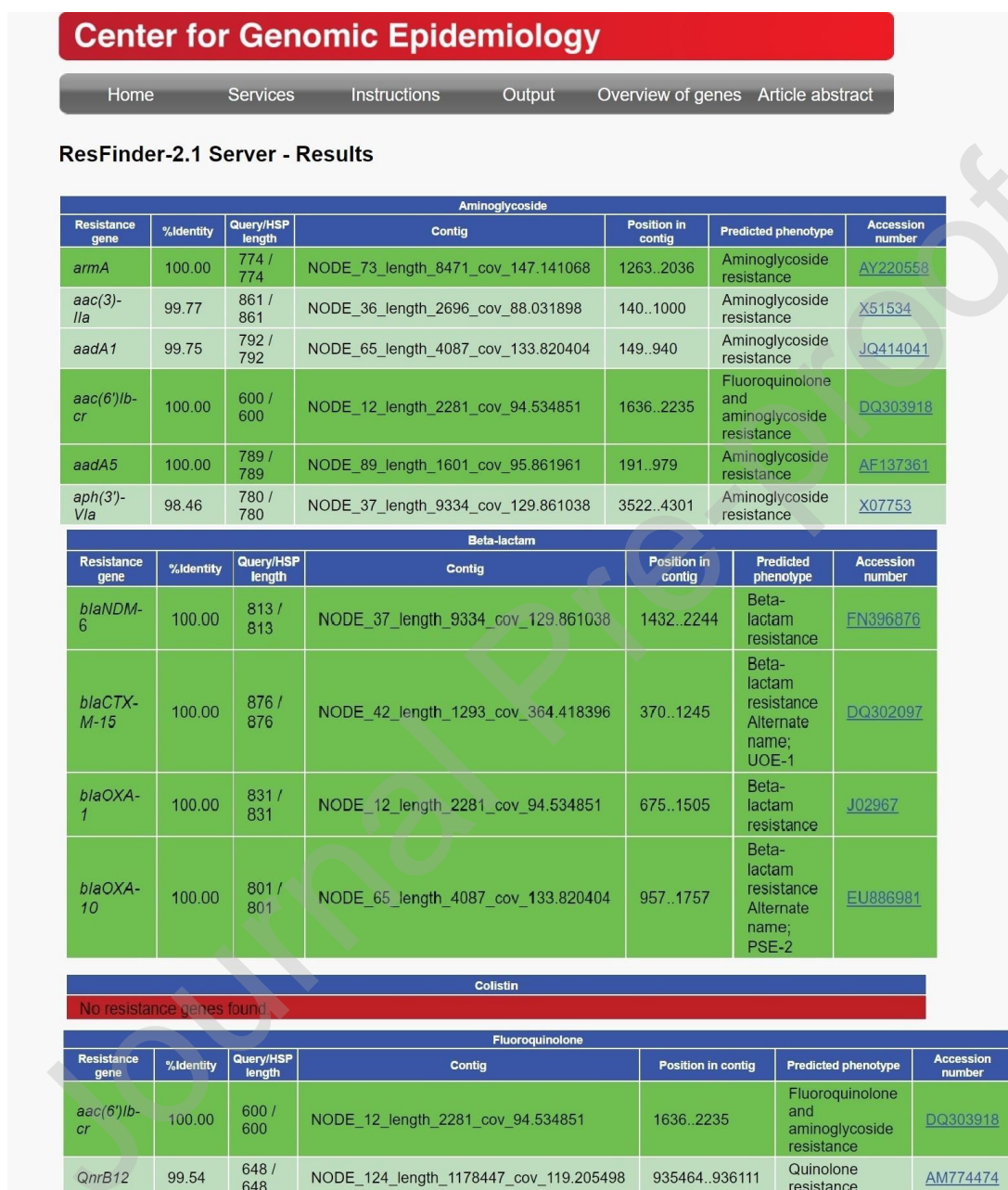
9. Kerr, M., Bray, B., Medcalf, J., et al. 2012. Estimating the financial cost of chronic kidney disease to the NHS in England. *Nephrol. Dial. Transplant.* 27:iii73-iii80.
10. Nayar, R., Shukla, I., Sultan, A. 2014. Epidemiology, Prevalence and Identification of *Citrobacter* Species in Clinical Specimens in a Tertiary Care Hospital in India. *IJSPR* 4:1-6.
11. Aziz, R.K., Bartels, D., Best, A.A., et al. 2008. The RAST server: rapid annotations using subsystems 119 technology. *BMC Genomics* 9:75.
12. Lowe, T.M., Eddy, S.R. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25:955–964.
13. Zankari, E., Hasman, H., Cosentino, S., et al. 2012. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* 67:2640–4.
14. Cosentino, S., Voldby Larsen, M., Møller Aarestrup, F., et al. 2013. PathogenFinder - Distinguishing Friend from Foe Using Bacterial Whole Genome Sequence Data. *PLoS One* 8:e77302.
15. Larsen, M.V., Cosentino, S., Rasmussen, S., et al. 2012. Multilocus Sequence Typing of Total-Genome-Sequenced Bacteria. *J. Clin. Microbiol.* 50:1355–1361.
16. Barakat, M., Ortet, P., Whitworth, D.E. 2013. P2RP: a web-based framework for the identification and analysis of regulatory proteins in prokaryotic genomes. *BMC Genomics* 14:269.
17. Zhu, H.Q., Hu, G.Q., Ouyang, Z.Q., et al. 2004. Accuracy improvement for identifying translation initiation sites in microbial genomes. *Bioinformatics* 20:3308–3317.
18. Ortet, P., De Luca, G., Whitworth, D.E., et al. 2012. P2TF: a comprehensive resource for analysis of prokaryotic transcription factors. *BMC Genomics* 13:628.
19. Ortet, P., Whitworth, D.E., Santaella, C., et al. 2015. P2CS: updates of the prokaryotic two-component systems database. *Nucleic Acids Res.* 43:D536–D541.

20. Finn, R.D., Tate, J., Mistry, J., et al. 2008. The Pfam protein families database. *Nucleic Acids Res.* 36:D281–D288.
21. Letunic, I., Doerks, T., Bork, P. 2012. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40:D302-5.
22. Tusnady, G.E., Simon, I. 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17:849–50.
23. Zojer, M., Schuster, L.N., Schulz, F., et al. 2017. Variant profiling of evolving prokaryotic populations. *PeerJ* 5:e2997.
24. Siguier, P., Perochon, J., Lestrade, L., et al. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* 34:D32–D36.
25. Brzuszkiewicz, E., Bruggemann, H., Liesegang, H., et al. 2006. How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proc. Natl. Acad. Sci. USA* 103:12879-84.
26. Roe, C.C., Vazquez, A.J., Esposito, E.P., Zarrilli, R., Sahl, J.W. 2019. Diversity, Virulence, and Antimicrobial Resistance in Isolates From the Newly Emerging *Klebsiella pneumoniae* ST101 Lineage. *Front. Microbiol.* 10:542. doi: 10.3389/fmicb.2019.00542.
27. Hannan, T.J., Totsika, M., Mansfield, K.J., et al. 2012. Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiol. Rev.* 36:616-48.
28. Barakat, M., Ortet, P., Jourlin-Castelli, C., et al. 2009. P2CS: a two-component system resource for prokaryotic signal transduction research. *BMC Genomics* 10:315.
29. Nishino, K., Nikaido, E., Yamaguchi, A. 2007. Regulation of multidrug efflux systems involved in multidrug and metal resistance of *Salmonella enterica* serovar Typhimurium. *J. Bacteriol.* 189:9066–9075.

30. Thomassin, J.L., Leclerc, J.M., Giannakopoulou, N., et al. 2016. Systematic analysis of two-component systems in *Citrobacter rodentium* reveals positive and negative roles in virulence. *Infect. Immun.* 85:pii:IAI.00654-16.
31. Altier, C., Suyemoto, M., Lawhon, S.D. 2000. Regulation of *Salmonella enterica* serovar Typhimurium invasion genes by *csrA*. *Infect. Immun.* 68:6790-6797.
32. Wei, B.L., Brun-Zinkernagel, A.M., Simecka, J.W., et al. 2001. Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. *Mol. Microbiol.* 40:245-256.
33. Suzuki, K., Wang, X., Weilbacher, T., et al. 2002. Regulatory circuitry of the CsrA/CsrB and BarA/UvrY systems of *Escherichia coli*. *J. Bacteriol.* 184:5130-5140.
34. Kirby, J.R. 2009. Chemotaxis-like regulatory systems: unique roles in diverse bacteria. *Annu. Rev. Microbiol.* 63:45–59.
35. Ulrich, L.E., Koonin, E.V., Zhulin, I.B. 2005. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 13:52-56.
36. Babu, M.M., Balaji, S., Aravind, L. 2007. General trends in the evolution of prokaryotic transcriptional regulatory networks. *Genome Dyn.* 3:66-80.
37. Minezaki, Y., Homma, K., Nishikawa, K. 2005. Genome-wide survey of transcription factors in prokaryotes reveals many bacteria-specific families not found in archaea. *DNA Res.* 12:269-280.
38. Troxell, B., Hassan, H.M. 2013. Transcriptional regulation by Ferric Uptake Regulator (Fur) in pathogenic bacteria. *Front. Cell. Infect. Microbiol.* 3:59.
39. Dong, T., Schellhorn, H.E. 2010. Role of RpoS in Virulence of Pathogens. *Infection and Immunity* 78:887-897.
40. Cascales, E., Buchanan, S.K., Duché, D., et al. 2007. Colicin biology. *Microbiol. Mol. Biol. Rev.* 71:158-229.

41. Skaar, E.P., Raffatellu, M. 2015. Metals in infectious diseases and nutritional immunity. *Metallomics*7:926-8.

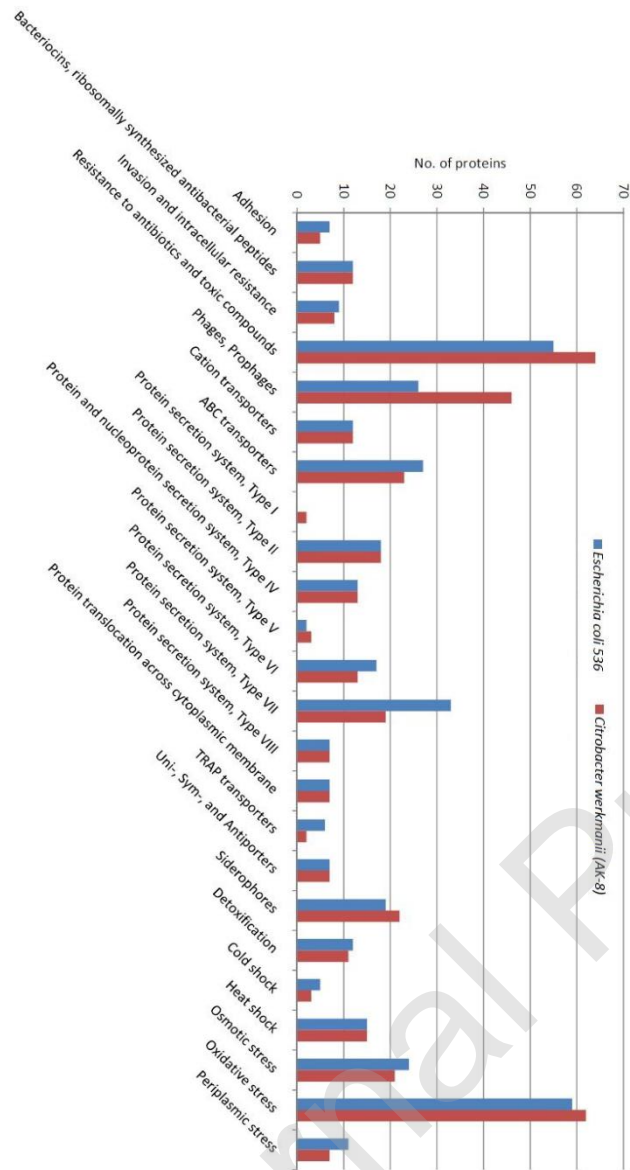
**Fig. 1A:** The screenshot of result of prediction of resistance determinants in *Citrobacter werkmanii* (AK-8) by ResFinder.





**Fig. 1B:** The screenshot of result of prediction of resistance determinants in *Citrobacter werkmanii* (AK-8) by ResFinder.

**Fig. 2:** Distribution of genes associated with different functional categories in *Citrobacter werkmanii* (AK-8) and *Escherichia coli* 356



**Table 1:** General genome features of *Escherichia coli* 536 and *Citrobacter werkmanii* AK-8

	<i>Escherichia coli</i> Strain 536	<i>Citrobacter werkmanii</i> Strain AK-8
Accession No.	NC_008253	NDXY000000000
Size (Mb)	4.94	5.22
Contigs	1	43
G+C	50.5	51.9
tRNA	80	61
rRNA	22	5
No. of RNAs	67	77
Genes	5,014	5,154
Proteins	4,623	5,077
Pseudo Genes	283	177
Scaffold	1	43
N <sub>50</sub>	4,938,920	3,76,299
L <sub>50</sub>	1	5
No. of subsystem	596	590
Coding sequences	4,753	4,900