

HHS Public Access

Diagn Microbiol Infect Dis. Author manuscript; available in PMC 2021 May 01.

Published in final edited form as:

Author manuscript

Diagn Microbiol Infect Dis. 2020 May ; 97(1): 114996. doi:10.1016/j.diagmicrobio.2020.114996.

Core Genome MLST and Resistome Analysis of Klebsiella pneumoniae Using a Clinically-Amenable Workflow

Madiha Fida1,* , **Scott A Cunningham**2,* , **Matthew P Murphy**2, **Robert A Bonomo**3,4,5, **Kristine M Hujer**3,4, **Andrea M Hujer**3,4, **Barry N Kreiswirth**6, **Nicholas Chia**7,8, **Patricio R Jeraldo**7,8, **Heidi Nelson**7,8, **Nicole M Zinsmaster**2, **Nikhil Toraskar**9, **Weizhong Chang**9, **Robin Patel**1,2 **Antibacterial Resistance Leadership Group**

¹Division of Infectious Diseases, Department of Medicine, Mayo Clinic, Rochester, MN

²Division of Clinical Microbiology, Mayo Clinic, Rochester, MN

³Louis Stokes Cleveland Department of Veteran Affairs Medical Center, Cleveland, OH

⁴Department of Medicine, Case Western Reserve University, Cleveland, OH

⁵Departments of Pharmacology, Biochemistry, Molecular Biology and Microbiology, and the Center for Proteomics and Bioinformatics, Case Western Reserve University, Cleveland, OH, and CWRU-Cleveland VAMC Center for Antimicrobial Resistance and Epidemiology (Case VA CARES), Cleveland, OH

⁶Public Health Research Institute, Rutgers University, Newark, NJ

⁷Center for Individualized Medicine, Mayo Clinic, Rochester, MN

⁸Department of Surgery, Mayo Clinic, Rochester, MN

⁹OpGen Inc, Gaithersburg, MD

Abstract

Whole genome sequencing is replacing traditional microbiological typing methods for investigation of outbreaks in clinical settings. Here, we used a clinical microbiology laboratory core genome multilocus sequence typing (cgMLST) workflow to analyze 40 isolates of K . pneumoniae which are part of the Antimicrobial Resistance Leadership Group (ARLG) isolate collection, alongside 10 Mayo Clinic K . pneumoniae isolates, comparing results to those of pulsed-field gel electrophoresis (PFGE). Additionally, we used the whole genome sequencing (WGS) data to predict phenotypic antimicrobial susceptibility (AST). 31/40 ARLG K. pneumoniae isolates belonged to the same PFGE type, all of which, alongside three isolates of different PFGE types, formed a large cluster by cgMLST. PFGE and cgMLST were completely

Correspondence address: Madiha Fida, MBBS, fida.madiha@mayo.edu, Division of Infectious Diseases, Mayo Clinic, 200 First Street SW, Rochester, MN 55905. *Both authors contributed equally

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. 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.

concordant for the 10 Mayo Clinic K. pneumoniae isolates. For AST prediction, the overall agreement between phenotypic AST and genotypic prediction was 95.6%.

Introduction

Klebsiella pneumoniae is a frequent pathogen in nosocomial settings, causing variety of infections, including hospital acquired pneumonia, bacteremia and urinary tract infection (1). An alarming increase in resistance to a broad array of antibiotics in this pathogen is now recognized, with several notable hospital-associated outbreaks reported (2–5). To definitively diagnose an outbreak, isolates need to be typed and shown to be genetically related. Various techniques are employed to type K . pneumoniae isolates to determine genetic relatedness. Now-unavailable methods like PCR electrospray ionization mass spectrometry (PCR/ESI-MS), alongside older methods such as pulsed-field gel electrophoresis (PFGE) and multi-locus sequence typing (MLST), are being replaced by whole genome sequence (WGS)-based typing methods. WGS data provides high-level differentiation amongst isolates of the same species and can also be used to predict phenotypic antimicrobial susceptibility (AST) (6, 7). WGS-based typing approaches rely on identification of single nucleotide polymorphisms (SNPs) across the entire genome, or genome-wide, gene-by-gene, allelic profiling of core genes or shared genes, known as core or whole genome multilocus sequence typing (cgMLST or wgMLST, respectively). cgMLST has been employed to differentiate outbreak and non-outbreak isolates of Acinetobacter baumannii, Listeria monocytogenes, Staphylococcus aureus, Legionella pneumophilia, Francisella tularensis, K. pneumoniae and Mycobacterium tuberculosis, among others (2, 8–14). Criteria for discrimination of outbreak from non-outbreak isolates (i.e., relatedness thresholds) are incompletely defined for SNP-based WGS analysis, cgMLST and wgMLST, and differ by species (8, 12, 15).

We described the use of cgMLST for clonality testing of S. aureus using an automated analytic system, SeqSphere+ (Ridom, Meunster, DE) (7, 16). Here, we applied the same computerized system to cgMLST analysis of previously characterized collections of K. pneumoniae, including outbreak-associated isolates, in comparison to PFGE. Additionally we used the WGS data to predict phenotypic resistance and susceptibility using the Acuitas® Whole Genome Sequence Analysis pipeline (OpGen, Inc., Gaithersburg, MD).

Materials and Methods

Bacterial isolates

Forty-two isolates from the ARLG and 10 from Mayo Clinic were studied (17). Two ARLG isolates were excluded; one failed to assemble after three sequencing attempts and another was a misidentified species - Escherichia coli - as shown by MALDI-TOF MS and phenotypic testing. ARLG isolates had been recovered from five institutions in the Eastern United States, and were previously described; they all harbored bla_{KPC} , of one of two types, 2 or 3 (Table 1) (17). Most of the isolates were from Mount Sinai Medical Center in New York (MSMCNY) (n=24) or Cleveland Clinic (CC) (n=9), with the rest being from University of Pittsburg Medical Center (UPMC) (n=3), University Hospitals Case Medical

Center (UHCMC) (n= 2) and the Louis Stokes Veterans Affairs Medical Center (LSVAMC) $(n=2)$.

cgMLST

DNA was extracted and purified using the Quick-DNA™ Fungal/Bacterial Miniprep Kit (Zymo Research Corp.) and normalized to 0.2 ng/ μ L with a Quantus[™] fluorometer (Promega, Madison, WI). Next generation sequencing libraries were prepared from normalized DNA with Nextera® XT (Illumina, San Diego, CA) and dual-indexed. Sequencing was performed on the MiSeq® platform (Illumina) using the V2 2 X 250 bp paired-end chemistry with a maximum pooling capacity of six libraries per flow cell, targeting a minimum coverage of 200X. Reads were processed for adapter and index cleaning utilizing the MiSeq reporter software in real-time. Adapter-clipped read files were imported into SeqSphere+ software version 4.1.9. Velvet de novo assembly and cgMLST were performed within the software suite following default settings. The SeqSphere+ embedded K. pneumoniae cgMLST schema was employed using K. pneumoniae genome NC_012731.1 as a reference genome derived from the analysis of 30 query genome sequences. Minimum spanning trees were generated from the final comparison data table. Grouping clusters obtained with cgMLST typing were compared to data generated by PFGE. In addition, MLST types were generated within the SeqSphere+ tool (18).

PFGE

ARLG isolates had been previously-subjected to PFGE (17). For the remaining 10 isolates, we performed PFGE according to the established method (19). Digestion of whole chromosomal DNA in agarose was performed with Xba I (Sigma-Aldrich St Louis, MO). The resulting restriction fragments were separated in CHEF Mapper® XA system (Bio-Rad Laboratories, Hercules, CA) with 1% agarose gel. The gels were then stained with ethidium bromide, illuminated under UV light and photographed in a Gel Doc™ XR+ Gel Documentation System (Bio-Rad).

Prediction of phenotypic AST

Susceptibility prediction were assessed using the 40 ARLG isolates. WGS assemblies were analyzed with the OpGen's Acuitas® Whole Genome Sequence Analysis pipeline, a highresolution tool which identifies antibiotic resistance genotypes using OpGen's Antibiotic Resistance Genes Database which consists of approximately 2,400 curated antibiotic resistance genes from the following databases: the Antibiotic Resistance Genes Database (ARDB), the Comprehensive Antibiotic Resistance Database (CARD), Resfinder and the Lahey Clinic (20–22). After antibiotic resistance genes were determined by querying assembled whole genome sequences of K , *pneumoniae* isolates against OpGen's Antibiotic Resistance Genes Database, a subset of antibiotic resistance genes were used as input for OpGen's Acuitas Lighthouse® prediction models to predict antibiotic resistance/ susceptibility phenotypes (23).

The isolates underwent phenotypic AST by agar dilution, with results interpreted according to CLSI guidelines (24). For the purposes of this study, isolates that tested susceptible by phenotypic AST were considered susceptible, as were those that were susceptible dose

dependent (SDD) to cefepime, and those that were either resistant or intermediate were considered non-susceptible.

Results

PFGE types

PFGE had previously classified the ARLG isolates into 21 unique types including A (which was further subgrouped into A1-A13) and non-A (B, C, D, F, G, H, which were further subgrouped into B1, B2, C1 and C2) groups (Table 1) (17). Among Mayo Clinic isolates, 4 were unique, with the rest grouping into one of two groups of three indistinguishable isolates each (Table 2).

MLST

Forty ARLG isolates were clustered into 8 MLST types and one indeterminate type. 7 MLST types (ST512, ST268, ST278, ST234, ST133, ST1697, ST34) were represented by single isolates, with the remaining 32 isolates (80%) belonging to ST258 (Table 1). Among Mayo Clinic isolates, 3 were ST231, 3 were ST101, 2 were ST14 and one each were ST147 and ST307 (Table 2).

cgMLST sequence types

The 40 ARLG isolates encompassed 25 unique cgMLST types with 37 unique allelic profiles. Thirty-four of the ARLG K. pneumoniae isolates formed a large cluster with a maximum cumulative distance (calculated by adding allelic differences between individual isolates) of 147 allelic differences between the most distant isolates by cgMLST. Six isolates were different from all isolates and from each other by 1796 allelic differences. For the ARLG isolates, four allelic relatedness thresholds were evaluated, $0, 6, 15$ and 50 allelic differences.One pair and one triplet of isolates were indistinguishable. Nineteen isolates were classified into five unique clusters based on a relatedness threshold of 6 allelic differences; 25 into 4 unique clusters based on a relatedness threshold of 15 allelic differences (with two pairs of clusters collapsing to one and one new cluster appearing compared to groupings with a relatedness threshold of 6 allelic differences); and all 44 isolates comprising one large cluster based on a relatedness threshold of 50 allelic differences (Figure 1). 21/24 isolates from MSMCNY clustered together in the same group when an allelic threshold of $\overline{50}$ was used whereas with an allelic threshold of $\overline{15}$, fifteen isolates were grouped into three groups. Similarly all but one of the CC isolates (ARLG-1712) grouped together using an allelic threshold of ≤50. Even after using the allelic threshold of 15, 7/9 isolates clustered in same group, suggesting clonal relatedness.

Of the 10 isolates obtained from Mayo Clinic, 6 grouped into 2 clusters, whereas 4 did not group into any specific cluster and differed from each other and the clustered isolates by 78 allelic differences (Supplemental Figure).

Correlation between PFGE, blaKPC-type, and cgMLST

There were four ARLG isolates with different PFGE types (D, F, G and H), that differed from each other and all other isolates by 1796 allelic differences, and had unique cgMLST

and MLST types. The two PFGE C type isolates (C1 and C2) differed from each other by 3,651 allelic differences; these isolates were from different institutions and of different bla_{KPC} types, suggesting that they were not related to one another. 31/40 ARLG K. pneumoniae isolates belonged to the same PFGE type (A with subtypes A1–13), all of which, alongside three isolates of different PFGE types, formed a large cluster with a maximum cumulative distance of 147 allelic differences (calculated by adding allelic differences between individual isolates) by cgMLST; this cluster (shown as Group 1 in Figure 1) included the four groups of isolates with ≤15 allelic differences. The three isolates in the large cluster of 34 isolates which were not PFGE type A included two which were PFGE type B (ARLG-1708, ARLG-1709) and one which was PFGE type G (ARLG-1743).

The two PFGE type B isolates differed from one another by 39 allelic differences, were from two institutions and of different bla_{KPC} -types suggesting that cgMLST better differentiated them than did PFGE. There were 4 PFGE A9 isolates (ARLG-1170, ARLG1725, ARLG-1739, ARLG-1740), all from the same institution and of the same bl a_{KPC} type, of which three were indistinguishable by cgMLST, with the fourth having 6 allelic differences from the other three; the single PFGE A11 type (ARLG-1727) had only one allelic difference from the cluster of three PFGE A9 isolates (ARLG-1170, ARLG-1725 and ARLG-1739) and was from the same institution, but of a different bla_{KPC} type. The single isolate of PFGE A12 (ARLG-1726) and one of the 9 PFGE A13 isolates (ARLG-1744) differed by a single allelic difference, and were from the same institution and of the same bla_{KPC} type. The other 8 PFGE A13 isolates differed from ARLG-1744 by 29 allelic differences, and from each other by 10 allelic differences; all but one were from the same instution and were all of the same bl_{KPC} type. The three PFGE A1 (ARLG-1168, ARLG-1736, ARLG- 1742) isolates were of the same bla_{KPC} type and from the same instution; they differed from one another by 18 allelic differences. The single PFGE A10 isolate (ARLG-1159) had 17 allelic differes from the most closely related isolate, which was the only PFGE A8 isolate (ARLG-1738) and was from a different institution and of the same bla_{KPC} type. There was one PFGE A7 isolate (ARLG-1730), and it differed from the most closely related isolate (ARLG-1738) by 22 allelic differences; the two were from the same instution and of the same bla_{KPC} type. There was a single PFGE A2 isolate (ARLG-1741) which differed by 15 alleles from the most closely related isolate, a PFGE A4 isolate (ARLG-1729) from the same institution and of the same bla_{KPC} type. ARLG-1729 (PFGE A4) was in turn 3 allelic differences away from another PFGE A4 isolate, ARLG-1737. However, the three other PFGE A4 isolates, ARLG-1728, ARLG-1745 and ARLG-1735 were more distantly separated by cgMLST; the last two differening from each other by 23 allelic differences and ARLG-1728 differing from ARLG-1729 by 34 allelic differences. There were two PFGE A5 isolates (ARLG-1163 and ARLG-1731) from two medical centers and of different bla_{KPC} types, that differed from one another by 48 allelic differences. Finally, there was one PFGE A6 isolate (ARLG-1713), which differed from the most closely related isolate (ARLG-1709, which was from a different institution but of the same bla_{KPC} type) by 33 allelic differences.

There was 100% concordance between PFGE and cgMLST groupings for Mayo Clinic isolates.

Correlation between phenotypic AST results and genetic prediction of phenotypic AST

Most isolates were resistant to multiple antibiotic classes. Susceptibility and nonsusceptibility predictions were made for ampicillin, cefazolin, ceftriaxone, ceftazidime, cefepime, ertapenem, meropenem, aztreonam, gentamicin, tobramycin, trimethoprimsulfamethoxazole (TMP-SMX), ciprofloxacin and levofloxacin. Table 3 summarizes the performance of WGS for prediction of phenotypic resistance in terms of the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and agreement of genotypic prediction with phenotypic AST. Agreement represents the correct prediction of phenotypic susceptibility or non-susceptibility. Overall agreement was 95.6% (confidence interval, 93.4– 97.2%). All isolates were resistant to cefazolin, ampicillin, ceftriaxone, ceftazidime, ertapenem, meropenem and aztreonam by phenotypic and genotypic AST. All, except one isolate, were resistant to cefepime. This isolate was dose dependent susceptible however was predicted to be resistant by genotypic testing. Overall agreement of phenotypic AST and genotypic prediction was >87% for all antibiotics, except gentamicin. For gentamicin there was a high false negativity rate (n=22, Table 4) with overall agreement of only 45%. In contrast, overall agreement for tobramycin was 100%.

β**-lactam resistance**

All isolates had two or more β-lactamase genes (Table 4). Thirty-three isolates had at least 4; four had 3, two had 2 and one had 7 β-lactamase genes. These genes included bla_{KPC-2} , bla_{KPC-3} , bla_{OXA-9} , bla_{OXA-2} , bla_{OKP-B} , bla_{SHV-1} , bla_{SHV-5} , bla_{SHV-11} , bla_{SHV-12} , bla_{SHV-26} , bla_{SHV-75}, bla_{SHV-167}, bla_{TEM-1}, bla_{TEM-1A}, bla_{TEM-118}, and bla_{TEM-192}. All 40 isolates were correctly found to be bla_{KPC} -positive, with the type correlating with what had been previously determined using PCR and amplified product sequencing (17).

All but five isolates (ARLG-1159, ARLG-1164, ARLG-1169, ARLG-1709 and ARLG-1728) harbored $bla_{\text{OXA-9}}$. The most common bla_{SHV} gene was $bla_{\text{SHV-11}}$, carried by 20 isolates, followed by $bla_{\text{SHV-167}}$ and $bla_{\text{SHV-12}}$ carried by 8 and 6 isolates, respectively. Twenty-two isolates had $bla_{\text{TEM-192}}$.

ARLG-1719 and ARLG-1721, which were indistinguishable by cgMLST, had identical resistance gene profiles, whereas, ARLG-1725, ARLG-1739 and ARLG-1170, which were also indistinguishable by cgMLST, had the same β -lactamase genes with the exception of their *bla*_{SHV} type.

Fluoroquinolone resistance

Fluoroquinolone resistance-associated efflux pumps, genes and gene mutations identified included *oqxA* and *oqxB*; *qnrA1* and *qnrB19*; and *gyrA* (S83I, and S83F), and *parC* (S80I), respectively. All isolates harbored $qqxA$ and $qqxB$ with ARLG-1165 and ARLG-1163 additionally harboring *qnrA1* and *qnrB19*, respectively. All isolates except four (ARLG-1165, ARLG-1169, ARLG-1171 and ARLG-1725) harboured a gyrA S83 mutation with gyrA S83I found in 34 isolates and gyrA S83F in two isolates (ARLG-1164 and ARLG-1712). Thirty four isolates carried parC S80I mutation as well.

All four isolates (ARLG-1165, ARLG-1169, ARLG-1171 and ARLG-1724) carrying the combination of *oqxA*, *oqxB*, without *gyrA* or *parC* mutations were susceptible to levofloxacin. All of these isolates, except ARLG-1724, were also susceptible to ciprofloxacin.

Aminoglycoside resistance

Aminoglycoside resitance genes detected included aac(3)-IVa, aac(6')-33, aac(6')-Ib, aacA4–8, aadA1, aadA2, aadA6, aadB, aph(3')-Ia, and aph(4)-Ia. One isolate had six aminoglycoside resistance conferring elements; 9 had five, 9 had four, 7 had three, 13 had two and 1 had one. There was wide variation in phenotypic AST for gentamicin. Despite having the same resistance genes identified, phenotypic susceptibility was different for ARLG-1168 (gentamicin -resistant) and ARLG-1736 (gentamicin-susceptible) (Table 4). Both isolates were resistant and genotypically predicted to be resistant to tobramycin. All eight isolates (ARLG-1167, ARLG-1710, ARLG-1711, ARLG-1718, ARLG-1719, ARLG-1720, ARLG-1721) carrying *aadA2* and *aacA4–8* had non-susceptible phenotypes to tobramycin and gentamicin but were genotypically predicted to be susceptible to gentamicin.

Trimethoprim-sulfamethoxazole (TMP-SMX) resistance

Several TMP resitance genes were identified, including *dfrA1, dfrA12, dfrA14* and *dfrA1*8; the most common was $dfA12$ which was present in 25 isolates, followed by $dfA14$, which was present in 13 isolates. Three isolates had *dfrA18*, with one having *dfrA1*. All isolates except ARLG-1709, ARLG-1731 and ARLG-1743 carried at least one TMP resistance element. Five isolates had two TMP resistance-associated genes.

Sulfonamide resistance genes identified included *sul1*, sul₂ and sul₂ carried by 32, 14 and 8 isolates, respectively. All isolates except ARLG-1169 had at least one sulfonamide resistance gene. Fifteen isolates had two sulfonamide resistance elements identified. All three isolates (ARLG-1731, ARLG-1164, ARLG-1743) that were susceptible to TMP-SMX carried resistance elements to either sulfonamide or trimethoprim but not both.

Discussion

Herein, we compared WGS analysis using cgMLST to PFGE for typing of K. pneumoniae and found a high degree of correlation between the two typing methods. Additionally, we used the same WGS data to predict phenotypic AST and found a high degree of correspondence between the two.

cgMLST can be highly discriminatory, though the ideal allelic threshold used to define relatedness has been incompletely defined. Our data provides information to assist in the interpretation of cgMLST of K. pneumoniae. For ARLG isolates, using an allelic threshold cutoff of 15 allelic differences, cgMLST identified four clusters of 2 to 15 isolates and, 15 isolates that differed from all four groups and each other by >15 allelic differences. In contrast, PFGE identified 21 unique subtypes including A (which was further subgrouped into A1-A13 containing 1 to 9 isolates) and non-A (B, C, D, F, G, H, which were further subgrouped into B1, B2, C1 and C2) groups. Using a higher allelic threshold cutoff of 50 , 34 isolates clustered into one group. At this higher threshold cutoff, results of cgMLST

classification broadly matched those of MLST; all but two isolates (ARLG-1735 and ARLG-1745) from this large cluster grouped into MLST type ST258. ARLG-1735 grouped into MLST ST512 and had 23 allelic differences from the closest member of ST258 (ARLG-1745); the MLST type for ARLG-1738 was indeterminate. ARLG-1743 grouped into type G by PFGE which was a non-A PFGE group; by cgMLST (using allelic threshold cutoff of 15), it clustered with other isolates of PFGE type A13. The MLST type of this isolate was ST258 which matched the cgMLST grouping. ARLG-1743 had 3,908 allelic differences from ARLG-1724 which was also grouped into G by PFGE. This discrepancy may be secondary to mislabeling at the Mayo Clinic, the ARLG laboratory or the origin laboratory.

Accurate typing is essential for outbreak investigations. Previously used typing methods for K. pneumoniae have included PFGE, PCR/ESI-MS, MLST, and amplified fragment-length polymorphism (AFLP) analysis (25–28). PFGE is one of the most common methods employed for typing of bacteria and is based on digestion of chromosomal DNA with infrequently-cutting restriction endonucleases. The resulting DNA fragments take a specific pattern of discrete bands on the gel. These patterns are then compared with one another to determine relatedness. The procedure itself is standardized; however, the interpretation is not always straightforward. MLST involves sequencing of a small number of housekeeping genes within a bacterial genome. Since it only involves a few loci, it lacks the discriminatory power needed to analyze outbreaks (but is useful for broadly categorizing strains across time and geographical distributions). To distinguish between different isolates, genome-wide analysis should ideally be executed. The resolution capacity of WGS-based methods such as cgMLST is higher than historical methods. And, once the data is generated, it can be shared and utilized by distant laboratories for data assessment, providing unambiguous data comparison. Inter-laboratory standardization of sequencing methods, coverage and typing schemes is evolving.

To date, several studies have investigated cgMLST or wgMLST in outbreak situations, including investigation of K , pneumoniae outbreaks. In a Dutch study, wgMLST was used to identify an outbreak of NDM-1 producing K . pneumoniae (2). A total of 162 K . pneumoniae isolates underwent WGS; MLST and wgMLST were then performed using the WGS data with SeqSphere+. In addition, resistome and plasmid analysis were performed using an online tool ResFinder. Twenty-six of 162 isolates clustered into one group with a maximum allelic difference of 5. Interestingly, the investigators compared these isolates with unrelated isolates originating from five other health care centers and found 7 isolates that clustered with the 26 outbreak isolates with a maximum allelic difference of 5. This cluster had a distance of 2,926 genes from its nearest neighbor. All 33 isolates belonged to same MLST type (ST873), carried the same plasmid and had small differences in their resistomes. In an outbreak of colistin-resistant K. pneumoniae carbapenemase (KPC)-producing K. pneumoniae in the Netherlands that involved two institutions (a hospital and a nursing home), wgMLST was employed (14). During this outbreak, six patients were found to be colonized with KPC-producing K . pneumoniae, and four enviromental samples were positive for KPC-producing K. pneumoniae. Isolates were typed using AFLP and MLST. WGS followed by MLST+ typing was performed using an in-house scheme with 3,102 K. pneumoniae genes as a reference in SeqSphere+. Out of the 3,102 target genes, 3,042 were

used in MLST+ typing as these were common to all six isolates. Three isolates of ten (6 patient and 4 environmental isolates), including two isolates from patients and one from the environment, were indistinguishable. Four, two and one isolate had 1, 5 and 2 allelic differences respectively from the three indistinguishable isolates. The maximum allelic difference between any two isolates was 10. MLST+ results in this study were found to be concordant with those of AFLP. All of these isolates belonged to the same MLST type, ST258. WGS data was also used for resistome analysis without making phenotypic predictions. In our study, the most common sequence type of K. pneumoniae was ST258, a hybrid sequence type deriving most of its genome (~80%) from ST11 (29). ST258 represents the majority of KPC-producing K , pneumoniae, has the ability to rapidly disseminate in healthcare settings, and is frequently associated with nosocomial outbreaks $(30-32)$.

WGS can also been used for resistome analysis (2, 14, 33). To take it one step further, and make it clinically-useful, this data can then be utilitized to predict phenotypic AST. Thus far, WGS data has been used to predict phenotypic resistance for A. baumannii, E. coli, K. pneumoniae, S. aureus, and M. tuberculosis, among others (34–39). WGS data has additionally been evaluated for prediction of MIC with reported accuracy up to 92% (37). Overall agreement of phenotypic AST and genetic prediction was 95.6% for our isolates, similar to what has been described in previous studies (23, 35). For gentamicin however, the overall agreement was only 45%. There was a high rate of false negativity (phenotypic resistance, WGS-predicted susceptibility). In contrast, the overall agreement was 100% for tobramycin. The reason for this discrepancy is not clear. In a previous study employing the same platform for genetic prediction of AST for K . pneumoniae, the agreement between phenotypic and genotypic AST for gentamicin was 86% (23). Aminoglycoside resistance is mediated by several mechanisms, including drug modifying enzymes, target site mutations, drug efflux pumps and porin mutations (40). Drug modifying enzymes differentially modify diverse aminoglycosides, based on the presence of target sites on particular aminoglycosides. A limitation of the approach applied is incomplete knowledge of all resistance genes and mutations. In addition, the relationship between genotype and phenotype is not linear. In our study isolates carrying similar genes had different phenotypic resitance profiles. Multiple mutations and/or resistance genes may work in a non-additive manner (antagonist or synergistic) to confer resistance and one mutation may confer resistance to multiple antibiotics. Also, phenotypic AST, which is the current gold standard, is not free from errors and may be operator or instrument dependent, in addition to depending on breakpoints applied. All ARLG isolates were highly resistant to β-lactam antibiotics, an important limitation, as addressed below. A mixture of phenotypes would allow better evaluation of performance of genotypic susceptibiltiy prediction.

We have described cgMLST typing for K . pneumoinae isolates, and have compared this typing method to the traditionally used typing method of PFGE. We have also described relatedness with different cgMLST allelic thresholds for K. pneumoniae; however this is also one of the main limitations of our study as further data are needed to precisely pinpoint allelic threshold criteria for determining relatedness. Previous studies considered isolates with 5–10 allelic differences as being related $(2, 14)$. Our report illustrates the potential of cgMLST as an epidemiological tool to replace PFGE and other traditional typing methods.

We have also used the WGS data to predict phenotypic antimicrobial susceptibility with an overall agreement of 95.6%, though high rates of resistance in the study isolates limited evaluation of specificity and negative predictive values.

In conclusion, we evaluated a discriminatory and easy-to-use cgMLST scheme for WGSbased typing for K. pneumoniae. We have been routinely performing cgMLST analyses in our clinical microbiology laboratory for over two and a half years and have found it to be a technologist-friendly methodology. Here, we show that cgMLST correlates with PFGE for K. pneumoniae, appearing to be more discriminatory. cgMLST is a suitable method for investigating the relatedness of K. *pneumoniae* isolates encountered in outbreak scenarios. Additional benefits of cgMLST are the ability to easily compare nucleotide sequence data between laboratories. We also show that WGS data can predict phenotypic AST.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank the Mayo Clinic Center for Individualized Medicine Microbiome Program for supporting this study. We thank Dag Harmsen for his thoughtful review of this manuscript. Dr. Patel is supported, in part, by UM1 AI104681. We thank the ARLG for providing isolates used in this study. We thank OpGen for providing the genotypic antimicrobial susceptibility prediction data. This study was also supported in part by funds the National Institute of Allergy and Infectious Diseases of the National Institutes of Health (NIH) to R.A.B. under Award Numbers R01AI100560, R01AI063517, and R01AI072219, and the Cleveland Department of Veterans Affairs Award Number 1I01BX001974 to R.A.B. from the Biomedical Laboratory Research & Development Service of the VA Office of Research and Development, and the Geriatric Research Education and Clinical Center VISN 10. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the Department of Veterans Affairs.

References

- 1. Podschun R, Ullmann U. 1998 Klebsiella spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin Microbiol Rev 11:589–603. [PubMed: 9767057]
- 2. Bosch T, Lutgens SPM, Hermans MHA, Wever PC, Schneeberger PM, Renders NHM, Leenders A, Kluytmans J, Schoffelen A, Notermans D, Witteveen S, Bathoorn E, Schouls LM. 2017 Outbreak of NDM-1-producing Klebsiella pneumoniae in a Dutch hospital, with interspecies transfer of the resistance plasmid and unexpected occurrence in unrelated health care centers. J Clin Microbiol 55:2380–2390. [PubMed: 28515215]
- 3. Woodford N, Tierno PM Jr., Young K, Tysall L, Palepou MF, Ward E, Painter RE, Suber DF, Shungu D, Silver LL, Inglima K, Kornblum J, Livermore DM. 2004 Outbreak of Klebsiella pneumoniae producing a new carbapenem-hydrolyzing class A beta-lactamase, KPC-3, in a New York Medical Center. Antimicrob Agents Chemother 48:4793–9. [PubMed: 15561858]
- 4. Chiang T, Mariano N, Urban C, Colon-Urban R, Grenner L, Eng RH, Huang D, Dholakia H, Rahal JJ. 2007 Identification of carbapenem-resistant Klebsiella pneumoniae harboring KPC enzymes in New Jersey. Microb Drug Resist 13:235–39. [PubMed: 18184049]
- 5. Samra Z, Ofir O, Lishtzinsky Y, Madar-Shapiro L, Bishara J. 2007 Outbreak of carbapenemresistant Klebsiella pneumoniae producing KPC-3 in a tertiary medical centre in Israel. Int J Antimicrob Agents 30:525–9. [PubMed: 17931835]
- 6. Junemann S, Sedlazeck FJ, Prior K, Albersmeier A, John U, Kalinowski J, Mellmann A, Goesmann A, von Haeseler A, Stoye J, Harmsen D. 2013 Updating benchtop sequencing performance comparison. Nat Biotechnol 31:294–6. [PubMed: 23563421]

- 7. Park KH, Greenwood-Quaintance KE, Uhl JR, Cunningham SA, Chia N, Jeraldo PR, Sampathkumar P, Nelson H, Patel R. 2017 Molecular epidemiology of Staphylococcus aureus bacteremia in a single large Minnesota medical center in 2015 as assessed using MLST, core genome MLST and spa typing. PLoS One 12:e0179003. [PubMed: 28575112]
- 8. Moran-Gilad J, Prior K, Yakunin E, Harrison TG, Underwood A, Lazarovitch T, Valinsky L, Luck C, Krux F, Agmon V, Grotto I, Harmsen D. 2015 Design and application of a core genome multilocus sequence typing scheme for investigation of Legionnaires' disease incidents. Euro Surveill 20.
- 9. Kohl TA, Diel R, Harmsen D, Rothganger J, Walter KM, Merker M, Weniger T, Niemann S. 2014 Whole-genome-based *Mycobacterium tuberculosis* surveillance: a standardized, portable, and expandable approach. J Clin Microbiol 52:2479–86. [PubMed: 24789177]
- 10. Antwerpen MH, Prior K, Mellmann A, Hoppner S, Splettstoesser WD, Harmsen D. 2015 Rapid high resolution genotyping of Francisella tularensis by whole genome sequence comparison of annotated genes ("MLST+"). PLoS One 10:e0123298. [PubMed: 25856198]
- 11. Leopold SR, Goering RV, Witten A, Harmsen D, Mellmann A. 2014 Bacterial whole-genome sequencing revisited: portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. J Clin Microbiol 52:2365–70. [PubMed: 24759713]
- 12. Higgins PG, Prior K, Harmsen D, Seifert H. 2017 Development and evaluation of a core genome multilocus typing scheme for whole-genome sequence-based typing of Acinetobacter baumannii. PLoS One 12:e0179228. [PubMed: 28594944]
- 13. Willems S, Kampmeier S, Bletz S, Kossow A, Kock R, Kipp F, Mellmann A. 2016 Whole-Genome Sequencing elucidates epidemiology of nosocomial clusters of Acinetobacter baumannii. J Clin Microbiol 54:2391–4. [PubMed: 27358465]
- 14. Weterings V, Zhou K, Rossen JW, van Stenis D, Thewessen E, Kluytmans J, Veenemans J. 2015 An outbreak of colistin-resistant Klebsiella pneumoniae carbapenemase-producing Klebsiella pneumoniae in the Netherlands (July to December 2013), with inter-institutional spread. Eur J Clin Microbiol Infect Dis 34:1647–55. [PubMed: 26067658]
- 15. Neumann B, Prior K, Bender JK, Harmsen D, Klare I, Fuchs S, Bethe A, Zuhlke D, Gohler A, Schwarz S, Schaffer K, Riedel K, Wieler LH, Werner G. 2019 A Core Genome Multilocus Sequence Typing Scheme for Enterococcus faecalis. J Clin Microbiol 57.
- 16. Cunningham SA, Chia N, Jeraldo PR, Quest DJ, Johnson JA, Boxrud DJ, Taylor AJ, Chen J, Jenkins GD, Drucker TM, Nelson H, Patel R. 2017 Comparison of whole-genome sequencing methods for analysis of three methicillin-resistant Staphylococcus aureus outbreaks. J Clin Microbiol 55:1946–1953. [PubMed: 28404677]
- 17. Endimiani A, Hujer AM, Perez F, Bethel CR, Hujer KM, Kroeger J, Oethinger M, Paterson DL, Adams MD, Jacobs MR, Diekema DJ, Hall GS, Jenkins SG, Rice LB, Tenover FC, Bonomo RA. 2009 Characterization of bla_{KPC}-containing Klebsiella pneumoniae isolates detected in different institutions in the Eastern USA. J Antimicrob Chemother 63:427–37. [PubMed: 19155227]
- 18. Jolley KA, Maiden MC. 2010 BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. [PubMed: 21143983]
- 19. MA P Chromosomal restriction fragment analysis by pulsed-field gel electrophoresis. Application to molecular epidemiology, p 651–7, Essential Procedures for Clinical Microbiology. ASM Press, Washington, DC.
- 20. Liu B, Pop M. 2009 ARDB--Antibiotic Resistance Genes Database. Nucleic Acids Res 37:D443– 7. [PubMed: 18832362]
- 21. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012 Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–4. [PubMed: 22782487]
- 22. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJ, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013 The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 57:3348–57. [PubMed: 23650175]

- 23. Walker GT, Quan J, Higgins SG, Toraskar N, Chang W, Saeed A, Sapiro V, Pitzer K, Whitfield N, Lopansri BK, Motyl M, Sahm D. 2019 Predicting Antibiotic Resistance in Gram-Negative Bacilli from Resistance Genes. Antimicrob Agents Chemother 63.
- 24. CLSI. 2019 Performance Standards for Antimicrobial Susceptibility Testing, 29th ed. Clinical and Laboratory Standards Institute, Wayne, PA.
- 25. Mammina C, Palma DM, Bonura C, Anna Plano MR, Monastero R, Sodano C, Cala C, Tetamo R. 2010 Outbreak of infection with *Klebsiella pneumoniae* sequence type 258 producing *Klebsiella* pneumoniae carbapenemase 3 in an intensive care unit in Italy. J Clin Microbiol 48:1506–7. [PubMed: 20200298]
- 26. van 't Veen A, van der Zee A, Nelson J, Speelberg B, Kluytmans JA, Buiting AG. 2005 Outbreak of infection with a multiresistant Klebsiella pneumoniae strain associated with contaminated roll boards in operating rooms. J Clin Microbiol 43:4961–7. [PubMed: 16207948]
- 27. Li J, Zou MX, Wang HC, Dou QY, Hu YM, Yan Q, Liu WE. 2016 An outbreak of infections caused by a Klebsiella pneumoniae ST11 clone coproducing Klebsiella pneumoniae carbapenemase-2 and RmtB in a Chinese teaching hospital. Chin Med J (Engl) 129:2033–9. [PubMed: 27569227]
- 28. Su LH, Leu HS, Chiu YP, Chia JH, Kuo AJ, Sun CF, Lin TY, Wu TL. 2000 Molecular investigation of two clusters of hospital-acquired bacteraemia caused by multi-resistant Klebsiella pneumoniae using pulsed-field gel electrophoresis and in frequent restriction site PCR. Infection Control Group. J Hosp Infect 46:110–7. [PubMed: 11049703]
- 29. Chen L, Mathema B, Pitout JD, DeLeo FR, Kreiswirth BN. 2014 Epidemic Klebsiella pneumoniae ST258 is a hybrid strain. MBio 5:e01355–14. [PubMed: 24961694]
- 30. Mezzatesta ML, Gona F, Caio C, Petrolito V, Sciortino D, Sciacca A, Santangelo C, Stefani S. 2011 Outbreak of KPC-3-producing, and colistin-resistant, Klebsiella pneumoniae infections in two Sicilian hospitals. Clin Microbiol Infect 17:1444–7. [PubMed: 21668577]
- 31. Tofteland S, Naseer U, Lislevand JH, Sundsfjord A, Samuelsen O. 2013 A long-term lowfrequency hospital outbreak of KPC-producing Klebsiella pneumoniae involving Intergenus plasmid diffusion and a persisting environmental reservoir. PLoS One 8:e59015. [PubMed: 23536849]
- 32. Mammina C, Bonura C, Di Bernardo F, Aleo A, Fasciana T, Sodano C, Saporito MA, Verde MS, Tetamo R, Palma DM. 2012 Ongoing spread of colistin-resistant Klebsiella pneumoniae in different wards of an acute general hospital, Italy, June to December 2011. Euro Surveill 17.
- 33. R DS, Pinto NA, Hwang I, Younjee H, Cho Y, Kim H, Yong D, Choi J, Lee K, Chong Y. 2017 Molecular epidemiology and resistome analysis of multidrug-resistant ST11 Klebsiella pneumoniae strain containing multiple copies of extended-spectrum beta-lactamase genes using whole-genome sequencing. New Microbiol 40:38–44. [PubMed: 28072891]
- 34. Kumburu HH, Sonda T, van Zwetselaar M, Leekitcharoenphon P, Lukjancenko O, Mmbaga BT, Alifrangis M, Lund O, Aarestrup FM, Kibiki GS. 2019 Using WGS to identify antibiotic resistance genes and predict antimicrobial resistance phenotypes in MDR Acinetobacter baumannii in Tanzania. J Antimicrob Chemother 74:1484–1493. [PubMed: 30843063]
- 35. Stoesser N, Batty EM, Eyre DW, Morgan M, Wyllie DH, Del Ojo Elias C, Johnson JR, Walker AS, Peto TE, Crook DW. 2013 Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J Antimicrob Chemother 68:2234–44. [PubMed: 23722448]
- 36. Bradley P, Gordon NC, Walker TM, Dunn L, Heys S, Huang B, Earle S, Pankhurst LJ, Anson L, de Cesare M, Piazza P, Votintseva AA, Golubchik T, Wilson DJ, Wyllie DH, Diel R, Niemann S, Feuerriegel S, Kohl TA, Ismail N, Omar SV, Smith EG, Buck D, McVean G, Walker AS, Peto TE, Crook DW, Iqbal Z. 2015 Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. Nat Commun 6:10063. [PubMed: 26686880]
- 37. Davis JJ, Boisvert S, Brettin T, Kenyon RW, Mao C, Olson R, Overbeek R, Santerre J, Shukla M, Wattam AR, Will R, Xia F, Stevens R. 2016 Antimicrobial Resistance Prediction in PATRIC and RAST. Sci Rep 6:27930. [PubMed: 27297683]

- 38. Nguyen M, Brettin T, Long SW, Musser JM, Olsen RJ, Olson R, Shukla M, Stevens RL, Xia F, Yoo H, Davis JJ. 2018 Developing an in silico minimum inhibitory concentration panel test for Klebsiella pneumoniae. Sci Rep 8:421. [PubMed: 29323230]
- 39. Long SW, Olsen RJ, Eagar TN, Beres SB, Zhao P, Davis JJ, Brettin T, Xia F, Musser JM. 2017 Population Genomic Analysis of 1,777 Extended-Spectrum Beta-Lactamase-Producing Klebsiella pneumoniae Isolates, Houston, Texas: Unexpected Abundance of Clonal Group 307. MBio 8.
- 40. Magnet S, Blanchard JS. 2005 Molecular insights into aminoglycoside action and resistance. Chem Rev 105:477–98. [PubMed: 15700953]

Figure 1. Minimum spanning tree showing core genome multilocus sequence type of the 40 Antibacterial Resistance Leadership Group *Klebsiella pneumoniae* **isolates.** Each isolate is designed with a unique number. Isolates in the same circle having no allelic differences between them. The lines between the circles show the numbers of allelic differences. Thirty-four isolates formed a large cluster (A) with a maximum cumulative distance of 147 allelic differences between the most distant isolates by cgMLST (colored gray), with the remaining isolates differing from the group and from each other by 1796 allelic differences.

CC, Cleveland Clinic; MSMCNY, Mount Sinai Medical Center in New York City; UPMC, University of Pittsburgh Medical Center; UHCMC, University Hospitals Case Medical Center; LSVAMC, Louis Stokes Veterans Affairs Medical Center

Table 1.

Antibacterial Resistance Leadership Group (ARLG) *Klebsiella pneumoniae* **isolates showing correlation between core genome multilocus sequence type (cgMLST) and pulsed-field gel electrophoresis type (17).**

The multilocus sequence type (MLST) shown was based on whole genome sequence assemblies configured from SeqSphere+ independently of cgMLST prediction. A minimum spanning tree illustrating the cgMLST data is shown in Figure 1.

MSMCNY, Mount Sinai Medical Center in New York City; UPMC, University of Pittsburgh Medical Center; UHCMC, University Hospitals Case Medical Center; CC, Cleveland Clinic; LSVAMC, Louis Stokes Veterans Affairs Medical Center

Table 2.

Mayo Clinic *Klebsiella pneumoniae* **isolates showing complete correlation between core genome multilocus sequence type (cgMLST) and pulsed-field gel electrophoresis type.**

The multilocus sequence type (MLST) shown was based on whole genome sequence assemblies configured from SeqSphere+ independently of cgMLST prediction. A minimum spanning tree illustrating the cgMLST data is shown in the Supplemental Figure.

Table 3:

Performance of genotypic resistance data for prediction of phenotypic antimicrobial resistance in Klebsiella pneumoniae.

PPV, positive predictive value; NPV, negative predictive value; CI, confidence intervals; N/A, not available

Table 4: Phenotypic antimicrobial susceptibility (AST) of *Klebsiella pneumoniae* **isolates.**

Gray blocks represent interpretative disagreement between phenotypic AST and genotypic susceptibility/nonsusceptibility prediction whereas white blocks represent interpretative agreement.

AMP, Ampicillin; CFZ, Cefazolin; CAZ, Ceftazidime; CRO, Ceftriaxone; FEP, Cefepime; MEM, Meropenem; ETP, Ertapenem; ATM, Aztreonam; CIP, Ciprofloxacin; LVX, Levofloxacin; SXT, Trimethoprim-sulfamethoxazole; GEN, Gentamicin; TOB, Tobramycin